

Dendro-Provenancing:

Beiträge zur Weiterentwicklung der dendrochronologischen Holzherkunftsbestimmung für historische und archäologische Fragestellungen

Abhandlung (kumulative Dissertation)
zur Erlangung der Doktorwürde
der Philosophischen Fakultät
der
Universität Zürich

vorgelegt von

Urs Gut

Angenommen im Frühjahrssemester 2020
auf Antrag der Promotionskommission bestehend aus

Prof. Dr. Philippe Della Casa (hauptverantwortliche Betreuungsperson)
Prof. Dr. Harald Bugmann
Prof. Dr. Adriano Boschetti

Zürich, 2020

Abstract

Dendro-Provenancing ist die dendrochronologische Methode zur Bestimmung des Wuchsortes eines Baumes anhand von Jahrringvariablen (meist des jährlichen Breitentzuwachses). Die Methode wird seit einigen Jahrzehnten regelmässig zur Beantwortung historisch-archäologischer Fragestellungen wie z. B. der Untersuchung von Holzhandelsnetzwerken oder der Klärung der Rohstoffherkunft angewendet. Das Ausmass, in welchem Wuchsmuster sich regional unterscheiden, ist für die Schweiz bislang nicht eingehend untersucht worden. Im Rahmen eines vom Schweizer Nationalfonds geförderten und in Kooperation von Lehrstühlen der Universität und ETH Zürich betreuten Dissertations-Projektes galt es, das Potential des Dendro-Provenancings für die Beantwortung historisch-archäologischer Fragestellungen im schweizerischen Alpen- und Voralpenraum auszuloten. Um die Durchführbarkeit zu garantieren und weil der dendrochronologische und historische Forschungsstand ausreichend fortgeschritten schienen, wurde das Projekt räumlich auf den Bereich der Holzversorgung der neuzeitlichen (1500-1850 n. Chr.) Stadt Zürich begrenzt. Es zeigte sich jedoch, dass die methodischen Grundlagen des Dendro-Provenancings nicht hinreichend ausgearbeitet waren, weshalb nicht mit der Untersuchung historischer Jahrringserien begonnen werden konnte. Nur an den Jahrringserien rezenter Bäume, für welche die Wuchsbedingungen genau beschrieben werden können, lässt sich untersuchen, ob sich regional charakteristische Wuchsmuster ausprägen. Folglich musste die methodische Grundlagenforschung bei heute lebenden Bäumen ansetzen.

In Form von drei Artikeln liegen nun die Ergebnisse dieser Forschung vor. Im ersten Artikel wurden die Grundannahmen des Dendro-Provenancings formuliert, das relevante Wuchssignal identifiziert und die Ausbildung lokaler Zuwachssignale anhand rezenter Waldstandorte im Untersuchungsgebiet evaluiert. Mit dem in Artikel I präsentierten statistischen Verfahren gelang es mit hoher Erfolgsquote zuvor anonymisierte Jahrringserien ihrem ursprünglichen Herkunftsbestand zuzuweisen. Im Rahmen eines weiteren Artikels (II) wurde der Frage nachgegangen, ob aus der gängigen dendrochronologischen Beprobungspraxis bei Bauuntersuchungen eine verzerrte Repräsentation des Zuwachssignals resultieren könnte. Der untersuchte, europaweite Datensatz zeigte keine Hinweise auf die Existenz von unterschiedlichen hochfrequenten Zuwachssignalen in Jahrringbreitenserien, die aus unterschiedlichen

Himmelsrichtungen an rezenten Bäumen entnommen worden waren. Somit besteht bislang kein Hinweis auf eine Verzerrung des fürs Dendro-Provenancing relevanten Signals aufgrund der etablierten Beprobungspraxis. Mit Artikel III gelang es schliesslich in einer Simulation das für die Herkunftsbestimmung entscheidende theoretische Konzept der Referenzchronologien zu untermauern. Vergleiche zwischen Jahrringserien unbekannter Herkunft und Referenzchronologien, die – so wird angenommen – das lokale Wuchssignal eines Gebiets repräsentieren, bilden die Grundlage für die räumliche Eingrenzung des Wuchsorates. Diese Annahme wurde bislang zwar mehrfach kritisiert, doch bestand noch kein Verfahren, das es erlaubte, das Risiko der Vermischung verschiedener lokaler Signale in Referenzchronologien zu evaluieren. Solche Signalvermischungen können bereits beim Aufbau von Gebäudechronologien auftreten, in denen das Holz aus verschiedenen Standorten verbaut wurde. Mit der Simulation in Artikel III wurde dargelegt, dass im Allgemeinen sehr hohe Anforderungen an den Umfang und die Qualität des dendrochronologischen Datensatzes erfüllt sein müssen, um die Etablierung lokaler Referenzen in einem Gebiet zu ermöglichen. Schliesslich verdeutlichte die Untersuchung des Fallbeispiels der Herkunftsbestimmung der neuzeitlichen Zuwachsreihen aus der Stadt Zürich, dass trotz Jahrzehntelanger, intensiver dendrochronologischer Tätigkeit diese Voraussetzungen im Untersuchungsgebiet nicht erfüllt sind. Dennoch gelang es zumindest die Höhenlagenherkunft einiger hundert Fichtenserien aus neuzeitlichen Gebäuden und Musikinstrumenten relativ zuverlässig einzugrenzen.

Nur gross angelegte Projekte werden es in Zukunft ermöglichen, Dendro-Provenancing in einem Gebiet auf belastbarer Datengrundlage durchzuführen. Zudem sind weitere methodische Entwicklungen zwingend notwendig, um die Zuverlässigkeit von Herkunftsbestimmungen zu erhöhen und deren allfällige Unzuverlässigkeit transparent kommunizieren zu können.

Dank

Dieses Dissertations-Projekt hätte ohne die Unterstützung zahlreicher Personen nicht durchgeführt werden können. Mein grösster, aber auch unbeschreiblicher Dank gilt Karin Gut. Es wäre eine eigene Abhandlung nötig, um ihre Unterstützung gebührend in Worte zu fassen.

Direkten Einfluss auf das Projekt hatten Niels Bleicher (fachlicher Berater), Philippe Della Casa (Hauptbetreuung) und Harald Bugmann (Ko-Betreuung). Ihnen danke ich für die zahlreichen Diskussionen, Kommentare und Hilfeleistungen sowie die akademische Freiheit, die mir bei der Durchführung dieses Projekts gelassen wurde.

Adriano Boschetti danke ich für die Mitwirkung in der Promotionskommission. Thomas Eissing für die Ermöglichung eines Gastsemesters in Bamberg sowie fachliche Diskussionen und Anregungen. Michael Thomaschett und Andreas Bräm für die Organisation dendrochronologischer Gebäudeuntersuchungen und den Zugang zu entsprechenden Datensätzen. Andreas Mäder für die Bereitstellung eines Arbeitsplatzes im dendrochronologischen Labor der Stadt Zürich.

Ronald Schmidt von Wildnispark Sihlwald sowie den jeweiligen Revierförstern im Untersuchungsgebiet möchte ich für die Beprobungserlaubnisse danken. Matthias Bolliger, Sirio Forabosco, Fritz Gut, Lukas Gut, Tryfon Jagoulis, Nicolai Lengacher, Julius Schmidt, Thomas Wäckerle, Felix Walder und Sarah Wicki danke ich für die geleistete Feldarbeit.

Nina Roth, Christoph Conradi, Konrad Knüsel und Yanhao Shi (ETHZ Statistik Beratung und Labor) danke ich für die Einführung in die Klassifikation von Zeitreihen und die Durchführung einer unpublizierten Studie (siehe Literaturverzeichnis Synopse) an einem Teil meiner Daten, die mich die methodischen Analogien zwischen klassischem Dendro-Provenancing und der weitverbreiteten Nearest Neighbor Klassifikation erkennen liessen.

Für das Gegenlesen bzw. das Lektorat der Synopse bedanke ich mich bei Lukas Emmenegger und Julius Schmidt.

Dieses Dissertations-Projekt (P0ZHP1_162299) wurde durch den Schweizerischen Nationalfonds zur Förderung der wissenschaftlichen Forschung finanziert.

Inhaltsverzeichnis

I.	Synopse	- 3 -
1	Einleitung.....	- 3 -
2	Bisherige Methodik und Forschungsgeschichte des Dendro-Provenancings ..	- 5 -
2.1	Kurze Einführung in relevante konzeptionelle Grundlagen der Dendrochronologie.....	- 5 -
2.2	Verortung des Dendro-Provenancings in den Grundkonzepten der Dendrochronologie.....	- 9 -
2.3	Forschungsgeschichte.....	- 11 -
3	Untersuchungsgebiet.....	- 14 -
4	Übergeordnete Fragestellungen der Dissertation	- 15 -
5	Methodische Beiträge der Dissertation.....	- 17 -
5.1	Zusammenfassung Artikel I	- 17 -
5.1.1	Hintergrund	- 17 -
5.1.2	Ergebnisse	- 18 -
5.1.3	Fazit für die Methodenentwicklung und die Herkunftsbestimmung von historischen Zuwachsserien	- 21 -
5.2	Zusammenfassung Artikel II	- 22 -
5.2.1	Hintergrund	- 22 -
5.2.2	Ergebnisse und Fazit für die Methodenentwicklung und die Herkunftsbestimmung von historischen Zuwachsserien	- 24 -
5.3	Zusammenfassung Artikel III	- 24 -
5.3.1	Hintergrund	- 24 -
5.3.2	Ergebnisse	- 26 -
5.3.2.1	Methodische Entwicklungen	- 26 -
5.3.2.2	Resultate	- 28 -
5.3.3	Fazit für die Methodenentwicklung und die Herkunftsbestimmung von historischen Zuwachsserien	- 30 -
6	Fallbeispiel: Holzherkunft der dendrochronologisch untersuchten, neuzeitlichen Gebäude der Stadt Zürich	- 31 -
6.1	Einführung	- 31 -
6.2	Dendrochronologische Datengrundlage	- 32 -
6.3	Methoden	- 36 -
6.4	Resultate	- 37 -
6.4.1	Ansatz I: Zwei höhenlagspezifische Initialchronologien	- 37 -
6.4.2	Ansatz II: 15 Bestandeschronologien als Initialchronologien	- 39 -
6.5	Diskussion und Schlussfolgerungen.....	- 44 -

7	Schlussbetrachtung	- 48 -
7.1	Synthese der methodischen Fortschritte und Resultate	- 48 -
7.2	Schlussfolgerungen für die weitere Methodenentwicklung	- 49 -
7.3	Schlussfolgerungen für die historische und archäologische Forschung ..	- 50 -
7.4	Schlusswort	- 52 -
8	Anhang	- 53 -
8.1	Tabellennachtrag zu Artikel I	- 53 -
8.2	Exkurs zur statistischen Modellierung der Höhenlagenherkunft anhand von Zuwachsvariablen	- 53 -
9	Literaturverzeichnis	- 59 -
II.	Appendix	- 69 -
A.	Eigenleistung	- 69 -
B.	Publikationen	- 70 -
(1)	Artikel I	- 70 -
(2)	Artikel II	- 86 -
(3)	Artikel III	- 96 -

I. Synopse

1 Einleitung

Holz war der universelle Werkstoff und primäre Energieträger der gesamten Menschheitsgeschichte (Winiwarter/Knoll, 2007, 58). Erst mit dem Beginn des Zeitalters fossiler Brenn- und chemisch synthetisierter Werkstoffe verlor Holz allmählich seine Vorrangstellung. Entsprechend bedeutend für die Kulturgeschichte sind historisch-archäologische Fragestellungen zu Holztransport und -handel (Meiggs, 1998). Ein prähistorisches Beispiel für diese Bedeutung ist die zentrale Frage der Holzversorgung der zum UNESCO Weltkulturerbe gehörenden, prähistorischen Feuchtboden- bzw. Seeufersiedlungen im zirkumalpinen Raum, die tausende von Holzproben lieferten (Billamboz, 2014; Bleicher, 2014; Gut, 2013). Während die prähistorische Versorgung in den zitierten Fällen eher lokal erfolgte, finden sich z. B. im Mittelmeerraum bereits früh historische Belege für einen bedeutenden, überregionalen Holzhandel (Meiggs, 1998; Kuniholm et al., 2007). Nur um ein berühmtes Beispiel herauszugreifen: Schon im Alten Reich (drittes Jahrtausend v. Chr.) wurde Zedernholz für den Schiffsbau vom Libanon ins heutige Ägypten verhandelt (Kuniholm et al., 2007; Rich et al., 2016). Weitere, zum Teil detaillierte Schriftzeugnisse zum Holzhandel finden sich im Zusammenhang mit dem Bau antiker Schiffsflotten wie z. B. derjenigen Athens, welches Ende des 5. Jh. v. Chr. und während des 4. Jh. v. Chr. seinen Bedarf zu grossen Teilen mit makedonischem Holz deckte (Meiggs, 1998, 123–132). In Mitteleuropa wurde der überregionale Holzhandel ab dem Mittelalter und spätestens ab der Neuzeit zur Deckung des wachsenden städtischen Holzbedarfes unverzichtbar (Eissing, 2010; Eissing/Dittmar, 2011; Grossmann, 1972). Die zeitweise kritische Versorgungslage führte mitunter zu gewaltsaufgeladenen Auseinandersetzungen: So kam es in Zürich im Jahr 1489 zu einem Volksaufstand, welcher massgeblich durch eine Änderung des Forstgesetzes angeheizt worden war. Das von Bürgermeister Hans Waldmann forcierte Gesetz sollte zur Erhaltung und Verbesserung der Gemeindewälder im Herrschaftsgebiet des Alten Zürichs beitragen, doch endete der Aufstand mit einem Regierungssturz und Waldmanns Hinrichtung (Weisz et al., 1983, 15–16).

Mithilfe der jahrringanalytischen Herkunftsbestimmung von Holz (Dendro-Provenancing) lassen sich historisch-archäologische Fragestellungen zu Holztrans-

port und -handel aus einer von den Schriftquellen unabhängigen Perspektive beleuchten. Dieses weite Anwendungspotential motivierte die methodischen Beiträge zur Weiterentwicklung des Dendro-Provenancings, welche den Kern der vorliegenden, kumulativen Dissertation bilden. So sehr die gegebenen historisch-archäologischen Beispiele zur Erforschung verlocken, die Weiterentwicklung der Methode kann weder bei prähistorischen noch bei historischen Hölzern beginnen. Es ist entscheidend, dass neu entwickelte Methoden an Holz bzw. Jahrringserien überprüft werden können, für welche die geographische Herkunft und die Standortsbedingungen zweifelsfrei bekannt sind. Andernfalls bleiben grundlegende Unklarheiten bestehen. Zum Beispiel könnten Jahrringserien aus weit voneinander entfernten Beständen, aber mit qualitativ ähnlichen Standortbedingungen, grössere Ähnlichkeiten aufweisen als Jahrringserien aus räumlich benachbarten Beständen mit unterschiedlichen Standortbedingungen. Eine solche Kontrolle über verschiedene Einflussfaktoren auf das Baumwachstum ist im Methodenentwicklungsprozess nur bei Jahrringmessungen an rezentem Holz möglich. Und auch dann nur, wenn dendrochronologische Proben direkt im Bestand an stehenden Bäumen entnommen werden. Folglich muss die hier vorgestellte Arbeit bei rezenten Hölzern beginnen.

Wie fast alle dendrochronologischen Methoden besitzt auch Dendro-Provenancing einen statistischen Methodenkern. Dendro-Provenancing wird heute hauptsächlich von Dendroarchäologen praktiziert und findet beispielsweise bei der Analyse von Holzhandelsnetzwerken oder der Herkunftsbestimmung von Schiffswracks Anwendung (Daly/Nymoen, 2008; Eissing/Dittmar, 2011). Deshalb wäre zu erwarten, dass die Auseinandersetzung mit den methodischen Grundlagen des Dendro-Provenancings gerade in der Dendroarchäologie ausgeprägt stattfände. Dies ist jedoch nur ansatzweise der Fall (Bridge, 2012; Daly, 2007a; Haneca et al., 2005). Entscheidende Grundlagen – etwa zur Einschätzung der Präzision von Herkunftsbestimmungen – fehlen oder sind noch in Entwicklung (Bridge/Fowler, 2019; Drake, 2018). Ebenfalls wurden die Annahmen der Methode bislang nicht präzise formuliert und es fehlt deren systematische Überprüfung an rezenten Hölzern. Diese Lücken motivierten die hier vorgestellten Beiträge, die mit allen technischen Details entweder bereits in internationalen, peer-reviewed Zeitschriften publiziert worden sind oder die zum Zeitpunkt der Abgabe zur Publikation eingereicht worden waren. Das Hauptaugenmerk der Synopse liegt deshalb nicht auf den Details der statistischen

Methoden, sondern auf der Kontextualisierung und der verständlichen Vermittlung der Resultate an interessierte Forschende in Geschichte und Archäologie.

Um den Einstieg zu erleichtern, wird zunächst ein Überblick zu den hier relevanten konzeptionellen Grundlagen der Dendrochronologie geboten. Mittels dieses Grundgerüsts wird anschliessend die bisherige Methodik des Dendro-Provenancings konzipiert und die Forschungsgeschichte in aller Kürze aufgerollt. Darauf folgt eine Darstellung der übergeordneten Fragestellungen und der daran anknüpfenden Teilaspekte der Dissertation, denen jeweils einer der oben genannten Artikel gewidmet wurde. Nach einer Aktualisierung des Methodenstandes wird an einem historischen Fallbeispiel aufgezeigt, welche Herausforderungen derzeit im Untersuchungsgebiet bestehen. Diese Diskussion mündet in ein abschliessendes Fazit zur bisherigen Anwendung der Methode in historischen und archäologischen Forschungsarbeiten. Zum Schluss werden mögliche Wege für die künftige Weiterentwicklung des Dendro-Provenancings skizziert.

2 Bisherige Methodik und Forschungsgeschichte des Dendro-Provenancings

2.1 Kurze Einführung in relevante konzeptionelle Grundlagen der Dendrochronologie

Das Baumwachstum variiert in Abhängigkeit von klimatischen, ökologischen und biotischen Faktoren (Fritts, 1976; King et al., 2013). Zur Veranschaulichung der prinzipiellen Einflussfaktoren auf die Varianz in Jahrringserien konzipierte Cook (1985) das „*linear aggregate model for tree-ring series*“:

Gesamtvarianz der Jahrringbreiten $G = \text{klimatisches Signal } C + \text{altersbedingter Wachstumstrend } A + \text{endogene Störungsimpulse } D_1 + \text{exogene Störungsimpulse } D_2 + \text{unerklärte, zufällige Wachstumsvarianz } E^1$

Dendrochronologische Datierungen basieren auf überregionalen Ähnlichkeiten im Baumwachstum, die klimatisch (bzw. durch das Wetter²) bedingt sind. Deshalb

¹ Die Annahmen der Linearität und Unabhängigkeit stellen grobe Vereinfachungen der realen Zusammenhänge dar. Zum Beispiel werden die verschiedenen Terme im Modell von Cook – ausser der zufälligen Varianz E – zudem durch die Genetik eines Baumes beeinflusst (King et al., 2013). Trotzdem bietet das Modell eine nützliche Grundlage zur Diskussion der hauptsächlichen Einflussfaktoren auf das Baumwachstum (Cook, 1985, 24).

werden verschiedene Methoden angewandt, um das klimatische Signal³ (C) in Jahrringserien zu verstärken. Bei rezenten Baumbeständen kann bereits durch eine gezielte Samplingstrategie ein bestimmtes Klimasignal verstärkt werden. Hat eine dendroklimatologische Studie beispielsweise die Rekonstruktion der Sommertemperatur vergangener Jahrhunderte anhand von Jahrringbreitenserien⁴ (Zuwachs- serien) zum Ziel, so werden z. B. nahe der Waldgrenze wachsende Bäume ausgewählt, für welche die Sommertemperaturen limitierend für den Jahrringbreitenzuwachs sind (Fritts, 1976, 17–19, 118–120; Schweingruber, 1983, 30–31). Für Jahrringserien aus historisch-archäologischem Kontexten ist es zumeist nicht möglich, das klimatische Signal durch gezielte Probenauswahl zu verstärken. Der ursprüngliche ökologische Standort ist bei Holz, das beispielsweise in Gebäuden verbaut wurde, in der Regel unbekannt. Es kann deshalb nur spekuliert werden, welche klimatischen Faktoren für das jährliche Breitenwachstum der verbauten Bäume limitierend gewesen sein mochten.

Unabhängig davon, ob gezieltes Sampling möglich ist, müssen die Rohwertserien der Jahrringbreiten statistisch aufbereitet werden, da das klimatische Signal von anderen, nichtklimatischen Störfaktoren überlagert wird (Schweingruber, 1983, 104–144). Besonders negativ auf spätere Korrelationsrechnungen zwischen Jahrring- serien wirkt sich der altersbedingte Wuchstrend (A) aus (Schweingruber, 1983, 88). Dieser resultiert aus der Tatsache, dass mit zunehmendem Alter ein immer ungünstigeres Verhältnis von Kronenoberfläche zu Stammdurchmesser entsteht (Schweingruber, 1996, 24–25). Deshalb werden die Jahrringe tendenziell schmäler, je älter ein Baum wird. Dieser sogenannte Alterstrend wird zumeist deterministisch als negativ-exponentieller Trend modelliert und mittels Trendbereinigung (Detrending) herausgerechnet (Cook, 1985, 37–39). Gängig ist dabei die Division der

² Der Begriff "Klima" ist normalerweise Langzeitbeobachtungen von meteorologischen Variablen (z. B. mittlere Temperatur oder mittlere Niederschlagssummen über ein 30-jähriges Zeitfenster) vorbehalten (WMO, 1989). Um Formulierungen zu vereinfachen, wird der Begriff "Klima" im Folgenden auch im Kontext kürzerer Beobachtungsperioden verwendet, welche eher in der Zeitspanne der Wettervariabilität liegen.

³ "Signal" bezeichnet im Kontext dieser Arbeit die von allen Individuen einer Population von Zeitreihen geteilte, gemeinsame Variabilität.

⁴ Es existieren verschiedene Variablen, wie Frühholz- und Spätholzbreite oder die Anzahl von Frühholzgefäßern, die mittels Jahrringserien erfasst werden können. Meistens wird jedoch das Signal in jährlichen Breitenzuwachsen (Zuwachsserien) untersucht.

Jahrringbreiten durch den modellierten Alterstrend, wodurch dimensionslose Jahrringindexserien entstehen (Cook, 1985, 37).

Exogene (D1) und endogene (D2) nichtklimatische Störungen des Baumbestandes können persistente Zuwachsreaktionen in Bäumen auslösen. Beispiele für exogene Bestandesstörungen sind Sturmschäden oder die Entnahme von Holz im Zuge der forstlichen Bewirtschaftung. Dagegen fallen unter endogene Störungen des Bestandes beispielsweise kleinräumige Bestandeslücken, deren Bildung auf das natürliche Absterben von Bäumen in Folge von Alter und/oder erhöhten Konkurrenzdrucks zurückzuführen ist. Exogene wie endogene Störungen wirken sich generell in kürzeren Zeitspannen (einige Jahre bis Jahrzehnte) auf das Baumwachstum aus als der altersbedingte Wuchstrend. Die daraus resultierenden Zuwachsreaktionen können aufgrund ihrer kürzeren Persistenz schlecht mit deterministischen Funktionen modelliert werden. Da die aus D1 und D2 resultierenden Veränderungen des Zuwachses das klimatische Signal C überlagern, müssen sie aber zur Verstärkung von C eliminiert werden (Cook, 1985, 28–39; Cook/Kairiukstis, 1990, 98–132; Schweingruber, 1996; Speer, 2010).

Dazu können Hochpassfilter und autoregressive Modelle an die bereits trendbereinigten Jahrringindexserien angepasst werden. Die Residualserien der autoregressiven Modellierung (bzw. der Output von Hochpassfiltern) sind frei von persistenten Zuwachsreaktionen, die aus D1 und D2 erwachsen (Cook, 1985, 50–51; Cook/Kairiukstis, 1990, 104–124). Diese Indexserien zweiten Grades widerspiegeln die Jahr-zu-Jahr Variabilität des Breitenwachstums eines Baumes und werden hier, in Abgrenzung zu den aus dem Detrending resultierenden Indexserien, als Residualserien bezeichnet. Die übriggebliebene Variabilität repräsentiert folglich die Reaktion des Baumes auf die jährliche Ausprägung klimatischer Bedingungen (C) sowie einen unerklärten, zufällig-individuellen Varianzanteil (E). Da bereits alle langfristigen Zuwachstrends und mittelfristigen Zuwachsreaktionen herausgerechnet wurden, wird im Zusammenhang mit der jährlichen Wachstumsvarianz von hochfrequenter Zuwachsvariabilität gesprochen (Schweingruber, 1983, 96). Dagegen ist der (multidekadal) Alterstrend (A) ein Beispiel für niederfrequente Zuwachsvariabilität; die mehrere Jahre bis einige Jahrzehnte andauernden Zuwachsreaktionen, die aus exogenen und endogenen Störungen (D1 und D2) erwachsen,

fallen in den Bereich der mittelfrequenten Zuwachsvariabilität (Cook, 1985, 28–39; Schweingruber, 1996).

Die unerklärte, zufällige Wachstumsvarianz (E) kann durch die Bildung von Chronologien (Begriffsklärung weiter unten) und die anschliessende Berechnung von Mittelwertserien (Mittelkurven) reduziert werden (Cook/Kairiukstis, 1990, 123–132). Die klimatisch bedingten, synchron verlaufenden Zuwachsschwankungen, die in allen Bäumen eines gegebenen Gebietes auftreten, können genutzt werden, um die Residualserien verschiedener Bäume in relative Synchronlage zueinander zu bringen (ist das absolute Endjahr – das letzte gemessene Jahr einer Probe – bekannt, entspricht dieses Vorgehen dem bekannten dendrochronologischen “Datieren”). Die daraus resultierende Schar von Residualserien wird als Chronologie bezeichnet (Cook/Kairiukstis, 1990, 137; Schweingruber, 1983, 80–101). Werden also beispielsweise die Residualserien von zehn Bäumen gleichen Alters in Synchronlage gebracht, so entsteht eine Schar von zehn Serien. Sind die Endjahre aller Serien identisch und die Bäume alle beispielsweise 50 Jahre alt, so entsteht eine in allen Jahren zehnfach belegte und 50 Wuchsjahre repräsentierende Chronologie. Die individuelle Wachstumsvarianz kann nun durch Berechnung der Mittelwertserie ausgemittelt bzw. eliminiert werden. Im Beispiel entstünde demnach eine Serie aus 50 Jahresmittelwerten.

Die dargestellte statistische Aufbereitung von Zuwachsseries führt zu einer Verstärkung der hochfrequenten, klimatisch bedingten Zuwachsvariabilität, die allen Bäumen einer Chronologie gemeinsam ist. Diese gemeinsame Variabilität wird als hochfrequentes, klimatisches Zuwachssignal bezeichnet (Cook, 1985; Cook/Kairiukstis, 1990). Diverse Studien haben gezeigt, dass sich klimatische Bedingungen auch auf den niederfrequenten Bereich der Zuwachsvariabilität auswirken (Diaz et al., 2011, 113–147). Oft fehlen für historisch-archäologische Jahrringserien jedoch die notwendigen Metadaten (etwa zu ökologischem Standort und Alterskohorte), um niederfrequente Klimasignale von anderen niederfrequenten Signalen, wie ähnlich verlaufenden Alterstrends, zu unterscheiden. Deshalb ist es in diesem Fall methodisch weniger verfänglich sich auf die Verstärkung des hochfrequenten Anteils von C zu beschränken (vgl. Kap. 2.2), besonders wenn eigentliche Klimarekonstruktionen – wie in der vorliegenden Arbeit – nicht das Ziel sind.

Für den Aufbau langer Chronologien zur Datierung historisch-archäologischer Jahrringserien ist eine zweifelsfreie Datierung jeder Einzelserie entscheidend. Die Datierung beruht auf einem statistischen Musterabgleich in allen denkbaren Synchronlagen (Endjahrpositionen) und der visuellen Kontrolle (Schweingruber, 1983, 82–87). Die bekanntesten statistischen Vergleichsrechnungen sind die Korrelationsrechnung mit anschliessender Berechnung des t-Wertes sowie die Gleichlängigkeitsrechnung (Bauch/Eckstein, 1970; Wigley et al., 1987). Beide Verfahren basieren auf der hochfrequenten (Jahr-zu-Jahr) Variabilität⁵. Trends und andere mittel- bis niederfrequente Varianzanteile müssen durch die statistische Datenvorbehandlung (Detrending und Hochpassfilterung, wie oben beschrieben) eliminiert werden, um Falschdatierungen zu vermeiden.

2.2 Verortung des Dendro-Provenancings in den Grundkonzepten der Dendrochronologie

Das hochfrequente, klimatische Signal ist die Grundlage aller bisherigen Dendro-Provenancing Ansätze. Vielfach werden von den erfassbaren Jahrringvariablen (je nach Holzart: Früh-/Spätholzbreite, Gefässfläche, Frühholzgefäßanzahl usw.) nur die Jahrringbreiten erfasst und untersucht, da sie besonders effizient zu messen sind (Bridge, 2012). Die meisten Studien basieren auf paarweisen Vergleichen zwischen Zuwachsseries unbekannter Herkunft und regionalen Referenzen (Mittelkurven von Referenzchronologien) (Eissing/Dittmar, 2011; Fraiture, 2009; Ważny, 2002). Von letzteren wird angenommen, dass sie lokale bis regionale Gemeinsamkeiten des Baumwachstums wiederspiegeln (Domínguez-Delmás et al., 2014). Eine Zuwachsserie unbekannter Herkunft wird mit allen zur Verfügung stehenden Referenzen verglichen. Der Vergleich erfolgt durch dieselben statistischen Vergleichsrechnungen, die auch zur Datierung verwendet werden (vergleiche Kap. 2.1). Allerdings wird beim Provenancing nicht die korrekte Synchronlage gesucht (diese wurde ja schon bestimmt), sondern es wird diejenige Referenzmittelkurve gesucht,

5 Die Gleichlängigkeitsrechnung etwa bezieht sich allein auf Übereinstimmungen im hochfrequenten Bereich der verglichenen Jahrringbreitenserien/Mittelkurven (Bauch/Eckstein, 1970). Aber auch bei den gängigen Verfahren der t-Wert Berechnungen wird nur das hochfrequente Signal berücksichtigt (Baillie/Pilcher, 1973; Hollstein, 1980; Wigley et al., 1987). Beim Verfahren nach Baillie/Pilcher (1973) wird zunächst ein gleitendes, 5-jähriges Mittel für jede der zu korrelierenden Jahrringbreitenserien/Mittelkurven berechnet. Diese Ausgleichskurve wird zur Indexierung beider Serien verwendet. Beim Verfahren nach Hollstein (1980) werden die ersten Differenzen berechnet und anschliessend logarithmisiert. Erst danach wird in beiden Verfahren der Pearson Korrelationskoeffizient sowie dessen t-Wert berechnet.

für welche die absolut höchsten Werte aus der Vergleichsrechnung resultieren (“*best matches*”). Diese Vergleichsrechnungen berücksichtigen – wie oben ausgeführt – allein hochfrequente Übereinstimmungen zwischen Referenz und Zuwachsserie. Dieses Vorgehen wird kaum je reflektiert, findet seine theoretische Begründung aber darin, dass Alterstrends und mittelfrequente Variabilität infolge ökologischer Prozesse zu allen Zeiten und in allen Regionen ähnlich auftreten können und daher wahrscheinlich keinen Beitrag zur Herkunftsanalyse leisten (Fritts, 1976; Schweingruber, 1996).

Oft ergeben die Vergleichsrechnungen mehrere hohe Werte für verschiedene regionale Referenzen, weshalb mehr oder weniger strikte Schwellenwerte definiert werden (Daly, 2007b). Diese legen fest, wie hoch beispielsweise ein t-Wert sein muss, damit dieser als “guter” Match gewertet wird. Sind schliesslich alle Vergleichswerte berechnet, werden die Werte entsprechend der geographischen Koordinaten der Referenzen kartiert (z. B. bei Eckstein/Wrobel, 2007). So entsteht eine räumliche Darstellung der berechneten Ähnlichkeiten zwischen Referenz und Zuwachsserie unbekannter Herkunft. Die Herkunftsregion wird schliesslich durch das Gebiet angeähert, in welchem “gute” Matches lokalisiert werden konnten (Daly/Nymoen, 2008).

Jüngst wurden für die Eingrenzung “guter” Matches neue Ansätze entwickelt (Boswijk/Fowler, 2019; Bridge/Fowler, 2019; Drake, 2018). Drake (2018) entwickelte ein Verfahren zum statistischen Testen von Herkunftshypothesen. Bridge/Fowler (2019) bzw. Boswijk/Fowler (2019) präsentierte eine Methode, bei welcher – vor Kartierung der paarweisen Vergleiche – das überregionale Wuchssignal (die Gesamtmittelkurve eines Untersuchungsgebiets) von den regionalen Signalen (Referenzchronologien) subtrahiert wird. Anschliessend werden sogenannte “*field correlations*” genutzt, um *best matches* zu lokalisieren. Trotz des intuitiv einleuchtenden Zwischenschrittes der Subtraktion des überregionalen Signals blieb bei Bridge/Fowler (2019) und Boswijk/Fowler (2019) jedoch eine starke Erhöhung der Präzision der Herkunftsbestimmungen aus. Des Weiteren basieren alle eben erwähnten Verfahren im Kern weiterhin auf dem Vergleich von Serien unbekannter Herkunft mit lokalen Referenzchronologien. Dass solche Referenzen lokale Wuchssignale repräsentieren, wird somit vorausgesetzt, wenn auch nicht ganz ohne kritische Reflektion dieses Konzepts (Bridge, 2000).

Anders als bei der Datierung stehen also die lokalen bis regionalen Unterschiede des klimatischen Signals und nicht die überregionalen Gemeinsamkeiten im Fokus des Dendro-Provenancings. Je unterschiedlicher die Mikroklimata und/oder Standortcharakteristika in einem gegebenen, geographisch-topographischen Raum, desto eher – so wird angenommen – unterscheiden sich die Referenzchronologien und -mittelkurven, die diesen Raum abdecken (Bridge, 2012; Haneca et al., 2005). Lokale, regionale und überregionale klimatische Phänomene überlagern sich jedoch (Diaz et al., 2011, 175–227). Folglich liegen in den Zeitreihen der Jahrringe auch nach der statistischen Vorbehandlung (Kap. 2.1) Mischsignale vor, die sich nicht scharf voneinander abgrenzen lassen. Die Unterscheidung lokaler, regionaler und überregionaler Wuchsmuster ist deshalb lediglich ein analytisches Hilfsmittel, welches die Beschreibung von signaltheoretisch mitunter weit komplexeren Zusammenhängen erleichtert.

2.3 Forschungsgeschichte

In den 1960er und 70er Jahren begannen Bauch und Eckstein mit dem Aufbau von Chronologien an niederländischen Eichenholzbildtafeln (Bauch/Eckstein, 1970). Da zahlreiche Kunstwerke des 17. Jh. mit vorhandenen, lokalen Chronologien nicht dendrochronologisch zu datieren waren, stellte sich sogleich die Frage nach der Herkunft der Holztafeln (Ważny, 2002).

Seither gehört die Fragestellung der Holzherkunft sowie Dendro-Provenancing zum Standardrepertoire dendrochronologischer Analysen. Nebst Untersuchungen an Kunstgegenständen sind solche an Gebäuden (Eissing, 2007; Sass-Klaassen et al., 2008) und Schiffen (Daly/Nymoen, 2008; Domínguez-Delmás et al., 2013; van Daalen/van der Beek, 2004) anzufügen. Ein detaillierter Überblick zu vergangenen Dendro-Provenancing Beiträgen findet sich bei Bridge (2012). In jüngerer Zeit sind zahlreiche weitere Studien erschienen (z. B. Domínguez-Delmás et al., 2018; Grabner et al., 2018; Rodríguez-Trobajo/Domínguez-Delmás, 2015), dennoch ist das von Bridge für Europa gezeichnete Bild über den Forschungsstand des Dendro-Provenancings noch weitgehend aktuell.

Die höchste Dichte an Dendro-Provenancing Studien besteht in Nord- und Westeuropa (Bridge, 2012; Fowler/Bridge, 2015); genauer im Bereich des bereits von Bauch und Eckstein (1970) untersuchten baltischen, v. a. neuzeitlichen Holzhandels. Dieser konzentrierte sich auf die Küsten und flössbaren Gewässer an und um die Ost- und Nordsee sowie auf die Britischen Inseln (Daly, 2011; Eckstein/Wrobel,

2007; Haneca et al., 2005; Ważny, 2002). Im selben Gebiet wurde auch die Holzherkunft zahlreicher, meist neuzeitlicher Schiffwracks untersucht (Daly, 2007a; van Daalen/van der Beek, 2004). Ein Hauptgrund für das Aufblühen von Dendro-Provenancing Studien im nord- und westeuropäischen Raum war die frühe Etablierung eines relativ engmaschigen Netzwerks von regionalen und lokalen Referenzchronologien (Bridge, 2012).

Eine hohe Chronologiedichte besteht ebenfalls seit langem in Mitteleuropa (Hollstein, 1980; Hughes, 1982, 105–158). Trotzdem setzte die Erforschung des Potentials des Dendro-Provenancings dort etwas später ein, nämlich in den 2000er Jahren (Eissing, 2007; Eissing/Dittmar, 2011, 2008). Auch in Südeuropa und im mediterranen Raum wurden in den letzten Jahrzehnten grosse Anstrengungen unternommen, um den Aufbau von Chronologien voranzutreiben (Domínguez-Delmás et al., 2015; Kuniholm et al., 2007; Ważny et al., 2014). Trotz vergleichsweise geringer Datengrundlage wurden auch in diesem Gebiet bereits zahlreiche Herkunftsuntersuchungen durchgeführt (Domínguez-Delmás et al., 2018; Kuniholm et al., 2007; Rich, 2013; Rodríguez-Trobajo/Domínguez-Delmás, 2015; Ważny, 2011).

Nicht in allen Gebieten Europas erzielte Dendro-Provenancing belastbare Herkunftsbestimmungen. Bridge beschrieb beispielsweise für die Britischen Inseln eine generell schwierigere Situation (Bridge, 2000). Er verglich die *best matches* von anonymisierten Zuwachsserien rezenter Eichen in einem Netzwerk rezenter Bestandeschronologien. Dabei entdeckte er einige problematische *matches* zwischen Zuwachsserien weit voneinander entfernter Bestände. Auch Clusteranalysen halfen in diesem Fall nicht, interpretierbare, räumliche Gruppierungen der untersuchten Bestände zu etablieren (Bridge, 2000). Folglich scheint die statistische Ähnlichkeit von Jahrringserien nicht zwingend eine Funktion der geographischen Distanz zu sein. Ein ähnliches Fazit wurde von Haneca et al. (2005) für Flandern gezogen.

Aufgrund dieses Umstandes fanden einige Verfeinerungen der Dendro-Provenancing Methode statt (Bridge/Fowler, 2019; Drake, 2018; Eissing/Dittmar, 2011; Jansma et al., 2014). Bereits im vorherigen Kapitel (2.2) wurden die Entwicklungen von Bridge/Fowler (2019), Boswijk/Fowler (2019) und Drake (2018) zur besseren Lokalisierung bester Matches bzw. objektiveren Prüfung von Herkunftshypothesen thematisiert. Nebst diesen ist besonders die Hinzunahme topographischer, geomorphologischer und bodenkundlicher Informationen zur Eingrenzung des potentiellen

Herkunftsgebietes zu erwähnen (Eissing/Dittmar, 2011; Jansma et al., 2014). Anhand solcher Informationen können unter Hinzuziehung historischer Quellen zu Holzhandel und -transport mittels Kartierungen bestimmte Herkunftsräume ausgeschlossen werden, wobei aber immer die Gefahr von Zirkelschlüssen besteht. In gewissen Gebieten erwies sich ebenfalls die Kartierung der potentiell natürlichen Vorkommen der untersuchten Baumarten als hilfreich für die Eingrenzung des Herkunftsgebietes (Domínguez-Delmás et al., 2013; Eissing/Dittmar, 2011).

Nebst Jahrringbreitenserien sind in den letzten Jahren vermehrt andere Jahrringvariablen auf ihr Dendro-Provenancing Potential hin überprüft worden. Erste Erfolge liessen sich jüngst beispielsweise durch die Erfassung von anatomischen Merkmalen wie Gefässflächen und -anzahl erzielen: Akhmetzyanov et al. (2019) gelang es anhand quantitativ erfasster, anatomischer Merkmale in Jahrringserien nordwestspanische Eichenstandorte in Süd-, Mittel- und Nordstandorte zu unterteilen. Dasselbe Gebiet liess sich zuvor aufgrund von Jahrringbreitenserien nicht unterteilen.

Auch Isotopenverhältnisse, wie diejenigen von $^{86}\text{Sr}/^{87}\text{Sr}$, können zur Herkunftsbestimmung genutzt werden (English et al., 2001; Hajj et al., 2017; Rich, 2013; Rich et al., 2016). In Chaco Canyon, New Mexico, gelang so die genauere Lokalisierung der Bauholzherkunft von Balken der sogenannten “great houses” (English et al., 2001). Allein anhand von Zuwachsserien war es dort zuvor nicht möglich gewesen zwischen verschiedenen, potentiellen Herkunftsgebieten zu unterscheiden. Weitere Untersuchungen existieren zu den Isotopenverhältnissen von $^{13}\text{C}/^{12}\text{C}$ bzw. der Nutzung von Pyrolyse in Kombination mit Gaschromatographie und Massenspektrometrie zur Definition eines molekularen Fingerabdruckes von Lignozellulose (Kagawa/Leavitt, 2010; Traoré et al., 2017). Zudem existieren erste Studien, in denen DNS zur Herkunftsbestimmung hinzugezogen wurde (Deguilloux et al., 2006; Rachmayanti et al., 2009). Der technische und finanzielle Aufwand solcher Untersuchungen lässt jedoch nicht erwarten, dass diese in naher Zukunft flächendeckend angewandt werden (Bridge, 2012). Zudem sind manche dieser Methoden nur unter gewissen Voraussetzungen erfolgversprechend. So kann beispielsweise im Fall der Strontium-Isotopen-Verhältnisse eine Herkunftsbestimmung nur durchgeführt werden, wenn sich die Isotopen-Verhältnisse potenzieller Herkunftsregionen hinreichend unterscheiden (English et al., 2001; Hajj et al., 2017).

3 Untersuchungsgebiet

Die Evaluation des Potentials von Dendro-Provenancing muss in der Schweiz mit der Untersuchung lebender Bäume beginnen (siehe Einleitung). Dementsprechend war ein geeignetes Untersuchungsgebiet zu definieren, in dem sich die rezenten Bestände möglichst nahe an historisch genutzten Wäldern befanden und für das bereits ein möglichst umfangreicher Datensatz historischer Zuwachsserien existierte. Beide Voraussetzungen waren im Einzugsgebiet des Holzhandels der neuzeitlichen Stadt Zürich am ehesten erfüllt (vgl. Kap. 6).

Die Holzversorgung der Stadt Zürich ist historisch sehr gut aufgearbeitet worden (Bitterli, 2009; Weisz et al., 1983). Schriftliche Quellen belegen den Holztransport über alle flössbaren Gewässer nach Zürich. Folglich lässt sich das Gebiet der Holzversorgung auf das hydrologische System der Flusstäler der Linth und Sihl sowie der grösseren Zuflüsse zum Zürich- und Walensee begrenzen (Artikel I, Fig. 1). Aufgrund der Durchführbarkeit konzentrierte sich die Beprobung rezenter Standorte jedoch auf die beiden nachweislich am häufigsten genutzten Wasserwege im Sihl- und Linttal (Weisz et al., 1983, 351–361). Ebenfalls umfangreich war der Holzhandel um den Zürichsee selbst (Weisz et al., 1983, 356–358). Deshalb wurde als Drittes das Oberseegebiet fokussiert untersucht.

Da der Aufbau eines genügend dichten Netzwerks rezenter Bestandeschronologien sehr arbeitsaufwendig ist und der Zeitrahmen der Dissertation begrenzt war, beschränkte sich die Beprobung auf die Baumart Fichte. Sie zählt zu den am häufigsten während der Neuzeit verbauten Holzarten und wächst zudem auf verschiedenen Höhenlagen im Untersuchungsgebiet (vgl. Weisz et al., 1983). Da bereits verschiedentlich höhenlagspezifische Wuchsunterschiede der Fichte festgestellt wurden (z. B.: Dittmar et al., 2012; Wilson/Hopfmüller, 2001), schien somit zumindest der Aufbau höhenlagspezifischer Chronologien realistisch.

Zur Beprobung wurden gezielt heutige Fichtenbestände in der Nähe historischer Holzeinschlagsgebiete ausgewählt (vgl. Artikel I, Introduction sowie Fig. 1), damit für eine allfällige spätere Erweiterung der rezenten Chronologien mittels historischer Zuwachsserien optimale Voraussetzungen gegeben waren (vgl. Kap. 6).

Die meisten bisherigen Dendro-Provenancing Studien fanden in geographischen Grossräumen mit – über weite Strecken – einheitlichem Klimaregime statt (vgl. Kap.

2.3, z. B.: Ważny, 2002; Daly, 2011; Eckstein/Wrobel, 2007; Fraiture, 2009). Im Vergleich zu diesen Arbeiten ist das Untersuchungsgebiet der Dissertation sehr klein. Es befindet sich jedoch im topographisch komplex gegliederten, nordöstlichen Voralpen- und Alpenraum, wo tendenziell ausgeprägtere Mikroklimata existieren, aber nur wenige Untersuchungen vorhanden sind (Eissing/Dittmar, 2011; Grabner et al., 2018). Im hier gewählten Untersuchungsgebiet könnten somit ausgeprägtere regionale Unterschiede im Baumwachstum bestehen, die für das Dendro-Provenancing vorteilhaft wären.

4 Übergeordnete Fragestellungen der Dissertation

Wie in der Forschungsgeschichte dargestellt wurde, sind nach wie vor Herkunftsbestimmungen vorherrschend, die auf Zuwachsserien basieren. Dies verwundert nicht, da bereits zur Datierung die Messung der Jahrringbreiten notwendig ist. Folglich entsteht kein zusätzlicher Aufwand für die Datenerhebung.

Es erstaunt allerdings, dass die Methode des Dendro-Provenancings anhand von Jahrringbreiten sich Jahrzehnte lang kaum gewandelt hat. Erst in jüngster Zeit sind einige methodische Neuerungen publiziert worden (Boswijk/Fowler, 2019; Bridge/Fowler, 2019; Drake, 2018). Auch diese Studien beruhen auf den eingangs dargestellten Prinzipien der Dendrochronologie (Kap. 2.1–2.2). Die regionale Ausprägung des klimatischen Signals C ist aber – wie bereits angeführt – keinesfalls leicht in theoretische Konzepte zu fassen. Zudem existieren nur relativ rudimentäre Versuche die Annahmen des Dendro-Provenancings auszuformulieren (Bridge, 2012; Daly, 2007a; Haneca et al., 2005). Letzteres ist jedoch von fundamentaler Bedeutung, da die klare Formulierung der impliziten Annahmen einer Methode es erst ermöglicht, deren empirische Grundlagen zu beurteilen.

Für das Dendro-Provenancing ist die Annahme des Zusammenhangs von statistischer und geographischer Ähnlichkeit entscheidend. Bislang existieren zwar einige Studien, die diesen Zusammenhang an rezenten Baumbeständen zu evaluieren suchten (Bridge, 2000; Bridge/Fowler, 2019; Daly, 2007a; Drake, 2018), jedoch wurden die Erkenntnisse aus solchen Studien oft nur unzureichend publiziert und sind kaum reproduzierbar (eine Ausnahme bildet die Studie von Drake, 2018). Es existieren bislang keine Standardmethoden, mit denen sich die räumliche Stabilität von regionalen oder lokalen Ähnlichkeiten im interindividuellen Baumwachstum

untersuchen und sich gleichzeitig das Potential allfälliger lokaler Ähnlichkeiten für die Herkunftsbestimmung ausloten lässt. Beispielsweise wird meist die Zusammengehörigkeit der Einzelserien, die etwa in eine lokale Gebäudechronologie einfließen, *a priori* angenommen, sobald bestimmte Schwellenwerte für den t-Wert erfüllt sind oder es andere Hinweise für die lokale Herkunft des beprobeden Materials (etwa Schriftquellen) gibt (Eissing/Dittmar, 2011; Domínguez-Delmás et al., 2014; Daly, 2007a). Ebenfalls wird oft davon ausgegangen, dass die Bildung lokaler Referenzchronologien (rezenten wie historischen) zur Verstärkung eines lokal- bzw. standortspezifischen Wachstumssignals führt (Eckstein/Wrobel, 2007; Eissing/Dittmar, 2011; Haneca et al., 2005). Ob aber sich abzeichnende Ähnlichkeitscluster zwischen Einzelserien mit Bestandeschronologien übereinstimmen, wurde eher selten untersucht (Bridge, 2000; Garcia-Gonzalez, 2008; Haneca et al., 2005; Savva et al., 2006).

Viel erstaunlicher noch: Die Frage, wie sich lokale bis regionale, klimatische Signale beim Aufbau von Referenzchronologien erhalten lassen, scheint bislang kaum im Zusammenhang mit Dendro-Provenancing diskutiert worden zu sein (Babst et al., 2018; Bridge, 2012; Diaz et al., 2011; Hollstein, 1980; Schweingruber, 1996; Speer, 2010). Bisherige Dendro-Provenancing-Ansätze gehen implizit davon aus, dass sich durch die Erweiterung rezenter Bestandeschronologien mit historischen Serien lokale, klimatische Signale erhalten lassen (vgl. Kap. 2.2). Zwar mag diese Annahme zunächst plausibel erscheinen, doch bei genauerer Betrachtung ist sie selten haltbar. Der Holzbedarf der Städte war in den wenigsten Fällen aus lokalen stadtdeigenen Besitzungen zu decken (Eissing, 2007; Grossmann, 1972). Folglich kam dem Holzhandel in den meisten Regionen eine grosse Bedeutung zu und das heute dendrochronologisch untersuchte Bauholz relativ klar definierbarer Altstädte stammt aus einer mehr oder minder grossen Bandbreite von Quellen, die teils weit auseinander liegen können (Eissing/Dittmar, 2011; Grossmann, 1972; Weisz et al., 1983). Die weit verbreitete Praxis, beispielsweise eine Mittelkurve für eine Altstadt zu berechnen, mag daher intuitiv naheliegend sein, ist aber häufig nur die Illusion der Repräsentation eines lokalen Signals. Die Kontamination der lokalen Referenzchronologien durch extern eingehandeltes Holz ist sehr wahrscheinlich. Deshalb muss untersucht werden, ob und mit welcher Robustizität das lokale bis regionale Zuwachssignal beim Aufbau von Referenzchronologien und -mittelkurven erhalten werden kann.

Diese teils massiven Lücken in den methodischen Grundlagen des Dendro-Provenancings waren ausschlaggebend für die Ausrichtung der vorliegenden Dissertation. Die spezifischen Fragestellungen der einzelnen Methodenbeiträge werden im folgenden Kapitel ausformuliert und in den Kontext der hier skizzierten, übergeordneten Fragestellungen gesetzt.

5 Methodische Beiträge der Dissertation

5.1 Zusammenfassung Artikel I

5.1.1 Hintergrund

Der paarweise statistische Vergleich der Ähnlichkeit von Zuwachsserien steht im Zentrum der Dendro-Provenancing Methode. Wie oben (Kap. 2 und 3) ausgeführt, ist die Annahme des direkten Zusammenhangs zwischen geographischer Nähe und statistischer Ähnlichkeit oft problematisch. Zudem werden diese sowie andere grundlegende Annahmen des Dendro-Provenancings selten präzise formuliert (Bridge, 2012; Daly, 2007a; Haneca et al., 2005). Deshalb wurden in der Einleitung zu Artikel I zunächst die Schlüsselannahmen des Dendro-Provenancings präzisiert:

1. Im Untersuchungsgebiet variiert das Baumwachstum ausreichend, was zur Ausprägung regional oder lokal charakteristischer Jahrringbreitenmuster führt.
2. Die (Un-)Ähnlichkeit im Zuwachs ist über statistische Ähnlichkeits- bzw. Distanzmasse quantifizierbar.
3. Die grösste statistische Ähnlichkeit wird zwischen Zuwachsserien festgestellt, die sich in nächster geographischer Nachbarschaft befinden.

Es existieren Studien an rezenten Beständen in einigen Regionen Europas (siehe Kap. 2.3). Für kleinräumige Gebiete im alpinen Raum jedoch ist unklar, wie stark statistische Ähnlichkeit und geographische Herkunft korrelieren. Bislang wurden hauptsächlich Clusteranalysen durchgeführt, um regionale Ausprägungen des Baumwachstums zu untersuchen (Bridge, 2000; Garcia-Gonzalez, 2008; Haneca et al., 2005; Savva et al., 2006). Dendro-Provenancing arbeitet aber nicht mit Klassifikationen, die aus Clusteranalysen hervorgehen, sondern (wie in Kap. 2.2 ausgeführt) mittels Klassifikation nach der höchsten statistischen Übereinstimmung mit Referenzchronologien (*best matches*).

Eine solche Klassifikation einer Zuwachsserie unbekannter Herkunft nach den *best matches* mit Referenzserien bekannter Herkunft entspricht dem statistischen

Verfahren der Klassifikation nach Nächsten Nachbarn (Nearest Neighbor Classification) (Abanda et al., 2019; Conradi et al., unpubliziert; Kaya/Gündüz-Öğüdücü, 2015; Schmitt, 2006; Xing et al., 2010). Der Nächste Nachbar einer zu klassifizierenden Zuwachsserie ist diejenige Referenzserie, die den höchsten Wert für eine statistische Vergleichsrechnung liefert. Zwar werden in der Literatur standardmäßig t-Werte nach Hollstein (1980) oder Baillie/Pilcher (1973) berechnet, doch es existiert gemäss Wissens des Autors der vorliegenden Synopse keine systematische Untersuchung, die wahrscheinlich machte, dass dies die einzigen und am besten fürs Dendro-Provenancing geeigneten Ähnlichkeitsmasse sind. Des Weiteren war für das Untersuchungsgebiet zu evaluieren, wie stark die Nächsten Nachbarn einer Zuwachsserie mit der räumlich-geographischen Nachbarschaft der entsprechenden Bäume zusammenhängen. Aus diesen Überlegungen ergaben sich die Untersuchungsziele von Artikel I:

1. Die Entwicklung eines statistischen Verfahrens zur Evaluation der drei Schlussannahmen des Dendro-Provenancings.
2. Die Untersuchung der Eignung verschiedener Ähnlichkeitsmasse für das Dendro-Provenancing.
3. Die Analyse allfälliger Klassifikationsfehler (falscher Herkunftsbestimmungen) mit dem Ziel das fürs Dendro-Provenancing relevante Zuwachssignal vor dem Hintergrund alpiner und voralpiner Umweltbedingungen besser zu verstehen.

Im Untersuchungsgebiet der Dissertation (Oberer Zürichsee, Linttal und Sihltal) wurden 15 rezente Fichtenbestände beprobt (4 bis 6 Bestände pro Region; Details siehe Artikel I, 2.1). Die Bestände deckten 3 Höhenlagen ab, nämlich: Hochlagen (>1500 m ü. M.), Mittellagen (1000–1500 m ü. M.) und Tieflagen (<1000 m ü. M.). Somit basierte die Untersuchung von Artikel I auf einer Datengrundlage von 401 Zuwachsserien.

5.1.2 Ergebnisse

Zur Überprüfung der Annahmen des Dendro-Provenancings wurde das Verfahren der “*k-Nearest Neighbors leave one-out cross-validation*” auf Basis des Ratings von “*scissor plots*” (im Folgenden abgekürzt als: k-NN) entwickelt (methodische Details zu *scissor plots* und k-NN siehe: Artikel I). Was hier als k-NN bezeichnet wird, unterscheidet sich von der verbreitet angewandten *k*-Nächsten Nachbarn Klassifikation dadurch, dass es sich nicht allein um eine Klassifikationsmethode, sondern primär um eine statistische Validationsmethode handelt. Bei der k-NN wird nicht eine

spezielle, optimale Einstellung für den k -Nächsten Nachbarn Klassifikator gesucht. Vielmehr soll das optimale Ähnlichkeitsmass identifiziert werden, welches über alle möglichen Klassifikatoren (alle Werte für k) die beste Performanz zeigt.

Das k-NN Verfahren basiert im Kern auf paarweisen Vergleichen zwischen Einzelzuwachsserien. Im Datensatz wird jeweils die Herkunft einer Zuwachsserie anonymisiert. Für alle anderen Zuwachsserien wird die Bestandesherkunft als bekannt angesehen. Diese Einzelserien mit bekannter Herkunft fungieren also als Referenzen für die Klassifikation bzw. Herkunftsbestimmung. Referenzchronologien werden beim k-NN nicht aufgebaut, denn dabei würde angenommen werden, dass die Einzelserien eines Bestandes bereits ein bestandesspezifisches Signal enthielten, das durch die Mittelkurvenbildung verstärkt würde. Doch dies ist genau eine der Annahmen des Dendro-Provenancings, die es zu überprüfen galt. Weiter wird beim k-NN Verfahren keine spezifische Anzahl Nächster Nachbarn festgelegt. Zum Beispiel könnte man eine Zuwachsserie lediglich anhand des einen Nächsten Nachbarn (*best match*) klassifizieren. Doch wie soll die Klassifikation ausfallen, wenn sich für den Vergleich einer anonymen Serie mit Referenzserien z. B. viele ähnlich hohe t-Werte finden lassen? Um die Willkürlichkeit der Festlegung der Anzahl zur Klassifikation konsulterter Nächster Nachbarn zu umgehen, wird beim k-NN Verfahren die Stabilität aller prinzipiell möglichen Klassifikationen untersucht. D.h. es wird nicht nur die Klassifikation nach dem einen Nächsten Nachbarn, sondern nach den zwei Nächsten Nachbarn, den drei Nächsten Nachbarn usw. berechnet und mittels *scissor plots* bestimmt, wie hoch und stabil die Klassifikations-Performanz insgesamt ist (methodische Details finden sich in Artikel I, 2.2 bis 2.4, als Anhang der Online-Version wurde auch der im Rahmen des Artikels programmierte R Code publiziert).

Die Effekte diverser statistischer Vorbehandlungsmethoden der Zuwachsserien (Detrending, Transformationen, autoregressive Modellierung und deren Kombinationen) sowie die Effekte der anschliessenden, unterschiedlichen Vergleichsrechnungen (Korrelation, Gleichhäufigkeit, Fraktionierte Gegenhäufigkeit) auf die Klassifikationsperformanz wurden mittels k-NN verglichen. Die jeweils unterschiedlichen Kombinationen von Vorbehandlungsmethoden und Vergleichsrechnungen wurden im Artikel als unterschiedliche Ähnlichkeitsmasse (“*proximity measures*”) bezeichnet.

Unter den Vorbehandlungsmethoden schnitten Hochpass-Filter bei diesem Vergleich generell und deutlich am besten ab (Details siehe Artikel I). Von den getesteten Vergleichsrechnungen erzielte das Verfahren der Berechnung von t-Werten für Korrelationskoeffizienten die besten Klassifikationsergebnisse.⁶ Ob dabei vor der t-Wert-Berechnung ein parametrischer Pearson's Korrelationskoeffizienten oder ein nicht-parametrischer Spearson's Korrelationskoeffizienten berechnet wurde, wirkte sich kaum auf die Performanz der k-NN aus.

Die Hochpass-Filterung (nur die in Artikel I als “*high-performance*” bezeichneten Methoden) der Zuwachsserien in Kombination mit der Berechnung von t-Werten erzielte einen mittleren Klassifikationserfolg von 71.8% bis 79.2% über die relevantesten Stufen der k-NN (“*open stages*”, siehe 2.4.2 in Artikel I). Das heisst, der grösste Teil der anonymisierten Zuwachsserien konnte ihrem Herkunftsbestand zugeordnet werden.

Klassifikationsfehler (falsche Herkunftsbestimmungen) waren nicht über alle Bestände gleich verteilt. Die meisten Fehler resultierten bei Zuwachsserien aus Fichtenbeständen, die sich auf 1000–1198 m ü. M. befanden. Relativ häufig wurden diese Zuwachsserien falschen Beständen der gleichen Höhenlage bzw. Beständen zugewiesen, die etwas tiefer lagen (829–1000 m ü. M.). Eine Teiluntersuchung zeigte schliesslich, dass die Differenz im Zuwachssignal zwischen einigen Beständen auf 829 bis 1198 m ü. M. geringer als auf anderen Höhenlagen war. Weiter wurde für den gesamten Datensatz ein abruper Anstieg von Klassifikationsfehlern festgestellt, wenn lediglich zwei Nächste Nachbarn zur Bestandeszuweisung verwendet wurden. Ansonsten jedoch ergaben sich nur geringfügige Unterschiede in der Klassifikationsperformanz zwischen den Beständen unterschiedlicher Regionen und Höhenstufen.

⁶ Entgegen der Einwände von Bridge/Fowler (2019), die sich im Zusammenhang von Dendro-Provenancing gegen die Berechnung von t-Werten aussprachen, war die k-NN Klassifikationsperformanz bei der Kombination von t-Wert Berechnung und unterschiedlichen Hochpass-Filtern im Datensatz von Artikel I durchgehend besser als bei der Kombination derselben Filter mit der blossen Berechnung von Korrelationskoeffizienten (wie von Bridge/Fowler (2019) vorgeschlagen). Aufgrund von Platzmangel waren diese Resultate mit Ausnahme der Ähnlichkeitsmasse t und r nicht in die publizierten Tabellen von Artikel I integriert worden. Die Resultate werden aber Aufgrund der Einwände von Bridge and Fowler (2019) nun nachgereicht (siehe: Anhang, Tab. A1). Die Berechnung kann mit dem im Supplementary Material von Artikel I veröffentlichten R-Code sowie den dort ebenfalls publizierten Daten nachvollzogen werden.

5.1.3 Fazit für die Methodenentwicklung und die Herkunftsbestimmung von historischen Zuwachsserien

Die Entwicklung der k-NN Methode ermöglichte einen konsistenten Vergleich unterschiedlicher Ähnlichkeitsmasse. Zwar existieren bereits einige Untersuchungen verschiedener dendrochronologischer Ähnlichkeitsmasse (Fowler et al., 2017; Wigley et al., 1987). Dabei stand jedoch die Eignung der betreffenden Masse für die Datierung im Vordergrund. Die vorliegende Untersuchung schliesst die bislang existierende Lücke einer analogen Untersuchung mit dem Fokus auf der Herkunftsbestimmung. Die Anzahl der in Artikel I verglichenen Ähnlichkeitsmasse ist im Vergleich zu früheren Cross-Dating Studien ausserordentlich hoch. Dadurch gelang es erstmals aufgrund einer soliden, öffentlich zugänglichen Datengrundlage zu belegen, dass für das Dendro-Provenancing das hochfrequente Zuwachssignal entscheidend ist. Zumindest im untersuchten Datensatz scheint die Zuwachsvariabilität im mittel- bis niederfrequenten Bereich zu wenig bestandesspezifisch ausgeprägt und deshalb für das Provenancing vernachlässigbar.

Aufgrund der hohen Klassifikationserfolgsrate mit der Kombination von Hochpass-Filtern als Vorbehandlungsmethode und der t-Wert-Berechnung als Vergleichsrechnungsmethode scheinen ausgeprägte, zwischenbeständliche Unterschiede im hochfrequenten Zuwachssignal im Untersuchungsgebiet zu bestehen. Angenommen, dass diese Signalunterschiede ebenfalls für historische Zeiten existieren, ist folglich eine der wichtigsten Voraussetzungen gegeben, dass Dendro-Provenancing prinzipiell auch an historischen Zuwachsserien durchgeführt werden kann.

Die erste Annahme des Dendro-Provenancings der Ausprägung regional oder lokal charakteristischer Jahrringbreitenmuster erscheint somit erfüllt. Da die statistische Herkunftsbestimmung vergleichsweise gut funktionierte, erhärtet sich auch die zweite Annahme des Dendro-Provenancings, dass die (Un-)Ähnlichkeit im Zuwachs über statistische Ähnlichkeits- bzw. Distanzmasse quantifizierbar ist. Die dritte Annahme, gemäss derer die grösste statistische Ähnlichkeit zwischen Zuwachsserien, die sich in nächster geographischer Nachbarschaft befinden, festgestellt wird, ist für den Grossteil des untersuchten Datensatzes zutreffend. Für die Bestände im problematischen Bereich der Höhenlage von 829 bis 1198 m ü. M. resultierten jedoch hohe statistische Ähnlichkeitswerte auch bei Vergleichen zwischen bestandesfremden Zuwachsserien mit nennenswerter Häufigkeit (siehe Artikel I, 3.).

Es ist hervorzuheben, dass die drei Annahmen des Dendro-Provenancings in keinem Untersuchungsgebiet als *a priori* angenommen werden dürfen. Zu häufig sind Untersuchungen mit gegenteiligen Ergebnissen publiziert worden (Bridge, 2000; Haneca et al., 2005). Ohne empirische Evaluation bewegen sich die genannten Grundannahmen, trotz ihrer theoretischen Plausibilität, im Bereich der Spekulation. Der in Artikel I präsentierte Ansatz der k-NN ermöglichte es, die grundlegenden Voraussetzungen des Dendro-Provenancings für das Untersuchungsgebiet zu evaluieren. Der methodische Ansatz ist zudem auf andere, zukünftige Untersuchungsgebiete übertragbar.

5.2 Zusammenfassung Artikel II

5.2.1 Hintergrund

Seit der Etablierung der Dendrochronologie ist bekannt, dass sich die Zuwachsserien verschiedener Räden im selben Stamm nicht vollständig gleichen. Daher wurde bereits bei Fritts (1976) nicht nur die bestandesweite Variabilität in den Zuwachsserien (der sogenannte „*between-tree noise*“) quantifiziert, sondern auch die Variabilität, die innerhalb einzelner Bäume auftritt (der sogenannte „*within-tree noise*“; Cook/Kairiukstis 1990). Um diese individuelle Variabilität zu fassen, werden in ökologischen Studien zumeist mehrere Räden pro Baum untersucht. Bei der Beprobung historischen Holzes ist dies jedoch häufig nicht möglich (siehe unten).

Jüngst hat eine Studie Hinweise für unterschiedliche Klima-Wachstumsbeziehungen in Bohrproben aus ein und demselben Stamm bei heute noch stehenden Bäumen geliefert, die in Abhängigkeit der Himmelsrichtung auftreten (Fang et al., 2015). Aus diesen unterschiedlichen Klima-Wachstumsbeziehungen könnten Unterschiede in den Wachstumssignalen beispielsweise zwischen Ostseiten-Chronologien und Südseiten-Chronologien erwachsen. Es ist weitgehend unerforscht, wie sich solche unterschiedlichen Klima-Wachstumsbeziehungen auf die Zuwachsvariabilität in unterschiedlichen Stammräden eines Baumes auswirken. Entsprechende Untersuchungen waren jedoch für die vorliegende Arbeit von Belang, weil es sich erstens streng genommen um einen extremen Fall von Dendro-Provenancing handeln würde, wenn man einzelne Bohrkerne gar noch innerhalb des Stammes lokalisieren könnte. Zum zweiten könnte dieses Phänomen auch direkte Auswirkungen auf die historische Holzherkunftsanalyse haben:

Bei der Beprobung historischer Holzobjekte kann nicht determiniert werden, in welcher Himmelsrichtung sich die Holzprobe im Baumstamm ursprünglich befunden hat. Zudem repräsentieren historische Zuwachsserien oft Einzelradien, weil die Holzobjekte schlecht zugänglich sind oder möglichst zerstörungsfrei Jahrringbreiten gemessen werden sollen. Falls grössere Zuwachsunterschiede zwischen den Radien eines Baumes existieren würden, könnten sie sich – bislang undetektiert – in den mittels dieser Zuwachsserien aufgebauten Referenzchronologien fortsetzen. Wären in bestimmten Referenzchronologien die Radien einer Himmelsrichtung übervertreten, könnte dies zu einer Verzerrung des Zuwachssignals der Referenzchronologie führen. Je nachdem, wie stark solche Unterschiede sich gestalteten, könnten Referenzchronologien entstehen, die anstatt eines lokalen Zuwachssignals ein himmelsrichtungsspezifisches Wuchssignal widerspiegeln. Solche Referenzchronologien wären für die Herkunftsbestimmung ungeeignet. Andererseits, falls das Bestandessignal dominant bliebe und dennoch ein himmelsrichtungsspezifisches Signal feststellbar wäre, könnte dieses Signal möglicherweise zur Rekonstruktion der ursprünglichen Orientierung einer historischen Zuwachsserie im Stamm genutzt werden.

Anhand von historischen Serien lässt sich – wie oben ausgeführt – die Hypothese himmelsrichtungsspezifischer Zuwachsvariabilität nicht überprüfen. Jedoch ergab sich anlässlich eines von Allan Buras (Technische Universität München / Universitäts- und Forschungszentrum Wageningen) organisierten und durch die *Association of Tree-Ring Research* unterstützten, dendrostatistischen Workshops die Möglichkeit diese Hypothese anhand eines europaweiten dendrochronologischen Datennetzwerkes rezenter Bestände zu untersuchen.

Der internationale Datensatz (10 Länder) umfasste 22 Bestände, die sich auf 8 verschiedene Baumarten verteilten, darunter 5 Fichtenbestände. Die Beprobung folgte einem einheitlichen Design. Es wurden je Bestand 15 Bäume beprobt. Pro Stamm wurden jeweils zwei Bohrkerne entnommen, einer von Süden und einer von Osten her. Bei den beprobten Standorten handelte es sich – bis auf wenige Ausnahmen – um geschlossene, gleichstufige und möglichst nur eine Baumart umfassende Bestände in flachem Gelände (Details siehe 2.1 Artikel II).

5.2.2 Ergebnisse und Fazit für die Methodenentwicklung und die Herkunftsbestimmung von historischen Zuwachsseries

Für keinen der Bestände liessen sich himmelsrichtungsspezifische Unterschiede in der hochfrequenten Zuwachsvariabilität zwischen Bohrkernen desselben Baumes feststellen (siehe Results Artikel II). Der himmelsrichtungsspezifische Einfluss von Umwelt und Klima auf die Zuwachsseries in einem Baum scheint zumindest für das hochfrequente Zuwachssignal vernachlässigbar. Methodisch betrachtet konnte damit gezeigt werden, dass der bauminterne Anteil der Variabilität für die Zwecke des Dendro-Provenancings vernachlässigt werden darf und damit die oftmals auf Einzeleradien basierende Beprobungspraxis der bauforschenden Dendrochronologie kein Problem für die spätere Auswertung darstellt. Streng genommen ist diese Aussage nur für solche Bestände zulässig, die den Untersuchten entsprechen. Dennoch bedeuten die Ergebnisse der Studie für das Dendro-Provenancing eine Entwarnung. Gleichzeitig existiert aber bislang auch keine Grundlage, die es erlaubt, die ursprüngliche Orientierung historischen Holzes zu bestimmen. Letztere Anwendung dürfte jedoch für Fragestellungen bezüglich historischen Holzhandels von sekundärem Interesse sein.

5.3 Zusammenfassung Artikel III

5.3.1 Hintergrund

In Artikel I wurde nachgewiesen, dass für das Dendro-Provenancing ausreichend grosse Zuwachsunterschiede zwischen rezenten Beständen im Untersuchungsgebiet vorkommen. In anderen Studien, die ähnliche Resultate anhand rezenter Bestände erzielt hatten, wurde schliesslich davon ausgegangen, dass sich die rezenten Zuwachsunterschiede beim Aufbau langer, historischer Chronologien erhalten (Bridge, 2000; Bridge/Fowler, 2019; Daly, 2007a; Drake, 2018). Es existiert jedoch nach Wissen des Autors der hier vorliegenden Arbeit keine Studie, welche diese fürs Dendro-Provenancing zentrale Annahme je eingehend untersucht hätte. Deshalb ist unklar, welche Vermischung von Bestandesignalen stattfinden kann, wenn eine rezente Bestandeschronologie mit Zuwachsseries aus historischen Gebäuden verlängert wird.

Folglich haben “lokale”, historische Referenzchronologien bis zu einem gewissen Grad Blackboxcharakter; zumindest was die Repräsentation lokaler oder regionaler Zuwachsvariabilität anbelangt. Offensichtlich wäre aber die Herkunftsanalyse anhand

von Chronologien mit einem Mischsignal mehrerer Regionen zum Scheitern verurteilt. Für die hier angestrebte, kleinräumige Bestimmung der Holzherkunft ist es folglich entscheidend, dass die Zuwachsvariabilität der Referenzchronologien ein weitgehend unvermischt und kleinräumig lokalisierbares Signal widerspiegelt. Wäre dies nicht der Fall, hätte das Dendro-Provenancing nur eine sehr grobe räumliche Auflösung. Folglich würde die Herkunftsbestimmung von der Konsultation historischer Quellen zum Holzhandel sowie von topographisch-hydrologischen Überlegungen abhängen. Das Dendro-Provenancing wäre, je nach Grobheit der räumlichen Auflösung, nicht imstande, zusätzliche Indizien für oder gegen eine auf historisch-topographischer Grundlage formulierte Herkunftshypothese zu liefern.

Das Ziel von Artikel III war deshalb abzuklären, ob und wie sich bestandesspezifische Zuwachssignale beim Aufbau von langen, historischen Referenzchronologien im Untersuchungsgebiet erhalten lassen. Gleichzeitig sollte quantifiziert werden, wie stark das Signal einer Referenzchronologie durch andere Signale kontaminiert werden darf, um trotzdem noch für Herkunftsbestimmungen verwendbar zu bleiben. Da eine solche Untersuchung an echten Referenzchronologien aufgrund der oben genannten Probleme nicht möglich ist, wurde ein Simulationsansatz gewählt. In einer Simulation können Zuwachsserien erzeugt werden, deren Bestandeszugehörigkeit bekannt ist. Folglich kann beim Chronologieaufbau überwacht werden, wie stark sich Zuwachsserien unterschiedlicher Bestandessherkunft in einer Referenzchronologie vermischen. Die Herausforderung bestand nun darin, eine in allen essentiellen Aspekten mit der Realität vergleichbare Situation zu simulieren (Maria, 1997). Dies bedeutet, dass ein Datensatz zu erzeugen ist, der alle relevanten Eigenschaften eines realen, dendrochronologischen Datensatzes besitzt. So wird von einem Simulationsmodell gefordert, dass die realen, sich überlagernden, bestandesspezifischen, regionalen und höhenstufenspezifischen Zuwachssignale möglichst genau abgebildet werden.

Der Simulationsansatz basiert somit auf zwei grundlegenden Konzepten: 1) Der Formulierung eines adäquaten Simulationsmodells, welches die wichtigsten, statistischen Eigenschaften der im Rahmen der Arbeiten zu Artikel I bereits erstellten Bestandeschronologien abbildet. 2) Der Generierung eines Pseudo-Datensatzes simulierter, historischer Zuwachsserien.

Die pseudo-historischen Zuwachsreihen sind Zufallsrealisationen der im Simulationsmodell definierten Pseudo-Bestandessignale. Etwas salopp und unscharf formuliert könnte man die pseudo-historischen Zuwachsreihen als Bohrproben aus einem Pseudo-Baum bezeichnen. Dieser würde in einem Pseudo-Bestand, in welchem das Wachstumssignal aller Bäume sich ähnelte. Wie hoch diese Ähnlichkeit bzw. wie stark das Pseudo-Bestandessignal ist, wurde im Simulationsmodell entsprechend der realen Verhältnisse in den beprobteten Beständen festgelegt.

Nachdem der Datensatz pseudo-historischer Zuwachsreihen mit Hilfe des Simulationsmodells generiert worden war, wurde die Erstellung von Pseudo-Referenzchronologien algorithmisch vollzogen. Der algorithmische Aufbau war dabei als Imitation der gängigen dendrochronologischen Praxis des Referenzchronologieaufbaus anhand von *best matches* (vgl. Kap. 2) zu konzipieren. Im Unterschied zu subjektiven, menschlichen Entscheidungen erforderte der algorithmische Chronologieaufbau eine strenge Definition der Entscheidungskriterien. Somit war der Chronologieaufbau reproduzierbar. Anhand der etablierten Referenzen galt es schliesslich den Grad der Durchmischung der verschiedenen Pseudo-Bestandessignale zu untersuchen. Die drei Hauptziele von Artikel III waren folglich:

1. Die Einführung und Evaluation eines Simulationsmodells zur Erstellung pseudo-historischer Zuwachschronologien.
2. Die Entwicklung eines Algorithmus', der den Prozess des Aufbaus von Referenzchronologien fürs Dendro-Provenancing nachahmt.
3. Die Untersuchung der Signalmischung in den Pseudo-Referenzchronologien sowie, daran anknüpfend, die Einschätzung des Potentials für die Etablierung echter Referenzchronologien mit erhaltenem regionalem oder lokalem Signal im Untersuchungsgebiet.

5.3.2 Ergebnisse

5.3.2.1 Methodische Entwicklungen

Da in Artikel I bereits nachgewiesen worden war, dass das hochfrequente Zuwachssignal für das Dendro-Provenancing entscheidend ist, beschränkte sich das Simulationsmodell ebenfalls auf dieses Signal. Mittels der mathematischen Methoden der Eigenwertdekomposition und linearen Transformation gelang es, 15 Pseudo-Bestandessignale zu generieren, die im Vergleich mit den originalen Bestandessmittelkurven eine praktisch identische Interkorrelationsstruktur aufwiesen. Das gewählte Simulationsmodell scheint deshalb adäquat die Ähnlichkeiten/Unähnlichkeiten

der originalen Bestandessignale abzubilden. Während das gemeinsame Intervall der realen Bestandeschronologien aber lediglich 63 Jahre abdeckte, deckten die simulierten Signale einen Zeitraum von 1000 Jahren ab. Aus diesen Pseudo-Bestandessignalen konnten in einem weiteren Schritt (Details siehe Artikel III) auch Pseudo-Einzelbaumserien bzw. pseudo-historische Zuwachsindexserien unterschiedlicher Längen erstellt werden.

Pro Pseudo-Bestand wurden solange pseudo-historische Zuwachsindexserien generiert, bis die Belegung der Pseudo-Bestandeschronologie in jedem Jahr 30 Serien betrug (Einzelserien-Ansatz). Da Dendro-Provenancing in der Regel aufgrund von Gebäude- oder Objektchronologien und nicht aufgrund von Einzelserien durchgeführt wird, wurde zudem ein weiteres Verfahren untersucht. Dazu wurden anstatt einer Pseudo-Zuwachsindexserie deren sechs generiert und zusammengemittelt. Das Resultat dieses Vorgangs waren sogenannte Pseudo-Objektchronologien (Objektchronologie-Ansatz). Es ist optimistisch anzunehmen, dass in der Realität immer alle sechs Pseudo-Zuwachsindexserien, die in eine Pseudo-Objektchronologie einfließen, aus demselben Bestand stammen. Deshalb wurden Simulationsvarianten durchgeführt, in welchen eine bzw. zwei Pseudo-Zuwachsindexserien eines fremden Bestandes in die jeweils generierten Pseudo-Objektchronologien einflossen.

In jedem Simulationssetting wurden so zwischen 6312 und 6518 Pseudo-Zuwachsindexserien bzw. Pseudo-Objektchronologien generiert. Damit der Aufbau von Pseudo-Referenzchronologien initiiert werden konnte, mussten zusätzlich für die ersten 150 Jahre 15 Initialreferenzchronologien erstellt werden.

Die algorithmische Etablierung von Pseudo-Referenzchronologien wurde mit der Einführung des *PREF-Constructor* Algorithmus gelöst. Vereinfacht ausgedrückt wies dieser Algorithmus jede Pseudo-Objektchronologie bzw. jede Pseudo-Zuwachsindexserie im Datensatz der jeweils statistisch ähnlichen Initialreferenzchronologie zu, wodurch diese sich verlängerte. Die statistische Ähnlichkeit wurde durch t-Werte für die paarweise Korrelation ausgedrückt. In den meisten bisherigen Studien wurden t-Werte ab einer bestimmten Höhe als "hoch" bezeichnet und – wenn auch nicht immer scharf in Zahlen ausgedrückt – Schwellenwerte für den t-Wert festgelegt (Bernabei/Bontadi, 2011; Daly, 2007b, 2007a; Eckstein/Wrobel, 2007; Rodríguez-Trobajo/Domínguez-Delmás, 2015; Ważny, 2002). Damit wird eine Bedingung für das Mindestmass an statistischer Ähnlichkeit gesetzt, die für einen Ver-

gleich zwischen einer Zuwachsserie unbekannter Herkunft und einer Referenzserie gefordert wird. Wie streng dieser Schwellenwert war, war meist der Erfahrung des Bearbeiters überlassen, wurde aber nicht empirisch hergeleitet. Um den Einfluss solcher Schwellenwerte auf die Komposition der mittels PREF-Constructor generierten Referenzchronologien zu untersuchen, wurden pro Simulationssetting (Simulation mit Pseudo-Zuwachsindexserie bzw. Pseudo-Objektchronologien) jeweils Durchgänge mit unterschiedlichen Schwellenwerten durchgeführt (nämlich: t-Wert ≥ 5 ; t-Wert ≥ 10 ; t-Wert ≥ 15 und t-Wert ≥ 20).

Da das Generieren der Pseudo-Zuwachsindexserien bzw. Pseudo-Objektchronologien als Zufallsziehung konzipiert wurde, musste die aus dieser statistischen Zufälligkeit erwachsende Variabilität ebenfalls untersucht werden. Dazu wurde jedes Simulationssetting 1000-mal wiederholt und die Variabilität der Signalzusammensetzung der generierten Pseudo-Referenzchronologien analysiert. Auf diese Weise liess sich ausschliessen, dass haltlose Aussagen anhand der Interpretation einzelner Zufallsrealisierungen gemacht werden.

5.3.2.2 Resultate

Im Fokus von Artikel III standen die Untersuchung der Signalmischung in den Pseudo-Referenzchronologien sowie die Einschätzung des Potentials für die Etablierung echter Referenzchronologien im Untersuchungsgebiet. Insbesondere wurde analysiert, wie schnell Pseudo-Referenzchronologien durch bestandesfremde Serien kontaminiert werden (Anteil bestandesfremder Serien $\geq 20\%$). Im Folgenden werden die Resultate lediglich summarisch dargestellt. Genaue Zahlen sind Artikel III zu entnehmen.

Ein niedriges Kontaminationsrisiko wurde nur für den undurchmischten Objektchronologie-Ansatz mit einem Schwellenwert von $t \geq 15$ festgestellt.⁷ Das heisst, solange bei der Erstellung von Pseudo-Objektchronologien keine bestandesfremden Serien eingemischt wurden, konnten mittels PREF-Constructor Algorithmus ungemischte, bestandesspezifische Pseudo-Referenzchronologien erstellt werden. Wurden jedoch bestandesfremde Pseudo-Zuwachsindexserien in die Pseudo-Objektchronologien

⁷ Zur groben Einordnung: Ein t-Wert von 5 wird bei der Datierung als relativ hoher Wert eingeschätzt und gilt oft als guter Hinweis für deren Korrektheit. Ein t-Wert von 20 hingegen kommt äusserst selten vor. Eine derart hohe Ähnlichkeit wird bisweilen gar als Indiz für Serien vom selben Baum gewertet (Hollstein, 1980; Wigley et al., 1987).

eingemischt, sanken Qualität (Belegungsdichte) und Länge der Pseudo-Referenzchronologien markant. In vielen Simulationsdurchgängen mit $t \geq 15$ war diese Abnahme so stark, dass für weniger spezifisch ausgeprägte Pseudo-Bestandessignale keine Referenzchronologien mehr aufgebaut werden konnten. Selbst im Simulationssetting ohne Durchmischung war im Median lediglich für ein Drittel aller generierten Serien die Herkunft bestimmbar, während der grösste Teil der Serien nicht zugewiesen, der kleinste Teil (ca. 1% der Serien) falsch klassifiziert wurde.

Eine weitere Erhöhung des Schwellenwertes auf $t \geq 20$ führte ebenfalls zu einem Einbruch der Klassifikationsrate (maximal 6% der Serien klassifiziert). Mit derart hohem Schwellenwert für die geforderte Mindestähnlichkeit waren kaum mehr gut belegte und lange Pseudo-Referenzchronologien zu etablieren.

Wurden die t-Wert-Schwellen für den Objektchronologie-Ansatz gesenkt, so stieg sukzessive die Rate falsch klassifizierter Serien. Bei einem Schwellenwert von $t \geq 10$ widerspiegelte noch ein relativ hoher Anteil der aufgebauten Pseudo-Referenzchronologien bestandesspezifische Signale. Im Median wurden etwa drei Fünftel der Serien korrekt klassifiziert, je ein Fünftel wurden aber falsch oder nicht klassifiziert. Zudem wurde bereits in einem nennenswerten Anteil (mindestens 5%) der Simulationsdurchgänge ein hoher Durchmischungsgrad in fast allen Pseudo-Referenzchronologien festgestellt.

Das Beimischen einer einzelnen bestandesfremden Serie in die Pseudo-Objektchronologien führte (bei einem Schwellenwert von $t \geq 10$) zu einer leichten Zunahme falscher Klassifikationen und einer leichten Abnahme der Belegungsdichte und des Belegungszeitraums der etablierten Pseudo-Referenzchronologien. Wurden jedoch zwei bestandesfremde Serien eingemischt, waren die Zunahme falscher Klassifikationen und die Abnahme der Belegungsdichte und des Belegungszeitraums markant.

Bei einem Schwellenwert von $t \geq 5$ mit dem Objektchronologie-Ansatz waren bereits im überwiegenden Teil der Simulationswiederholungen die meisten Referenzchronologien stark durchmischt (Kontamination zum Teil weit über 20%).

Noch weniger zuverlässig erhielten sich bestandesspezifische Signale mit dem Einzelserien-Ansatz. Mit letzterem war es allein mit dem niedrigen Schwellenwert

von $t \geq 5$ möglich, ausreichend belegte und lange Referenzchronologien für die weitere Analyse zu etablieren. Bei höheren Schwellenwerten konnten für die meisten Pseudo-Bestandessignale mit dem Einzelserien-Ansatz keine Pseudo-Referenzchronologien mehr aufgebaut werden.

Ausser dem Erhalt bestandesspezifischer Signale wurde ebenfalls untersucht, ob sich regionsspezifische oder höhenlagspezifische Signalunterschiede in den Pseudo-Referenzchronologien erhalten lassen. Von den letzteren Zweien erwiesen sich allein höhenlagspezifische Signalunterschiede als robust. Selbst mit dem Einzelserien-Ansatz und einem Schwellenwert von $t \geq 5$ mischten sich Serien aus Hochlagen nicht mit solchen aus Tief- oder Mittellagen. Die letzten beiden Lagen konnten allerdings nicht zuverlässig voneinander getrennt werden.

5.3.3 Fazit für die Methodenentwicklung und die Herkunftsbestimmung von historischen Zuwachsserien

Die Simulation hat gezeigt, dass der Aufbau bestandesspezifischer Referenzchronologien auch für reale, historische Zuwachsserien prinzipiell möglich ist. Dieser Erfolg ist für das Dendro-Provenancing von grosser Bedeutung, da somit erstmals mittels einer intersubjektiv nachvollziehbaren Methode Belege für die konzeptionelle Richtigkeit des Referenzchronologieaufbaus angeführt werden können. Jedoch sind die erforderlichen Schwellenwerte extrem hoch ($t \geq 15$). Ebenfalls kann für reale Objektchronologien – etwa aus historischen Gebäuden – kaum eine rein lokale Holzherkunft angenommen werden. Zumindes gelegentlich werden auch in Gebäuden, für deren Bau lokale Ressourcen verwendet wurden, nicht lokale Hölzer verbaut worden sein. Wie gezeigt wurde, ist der Aufbau von bestandespezifischen Chronologien – und damit einhergehend auch der Aufbau regionaler Referenzchronologien – in der Simulation nur mit undurchmischten Objektchronologien möglich. Das Beimischen nicht-lokalen Serien verhinderte in der Simulation im Extremfall den Aufbau von Referenzen. Für reale Untersuchungen bedeutet das, dass die Annahme der lokalen Holzherkunft bei der Auswahl von Proben aus historischen Gebäuden und anderen Objekten sehr kritisch reflektiert werden muss. Beim Erstellen von Objektchronologien sollten – auch wenn die lokale Holzherkunft aufgrund der Topographie und/oder Schriftquellen wahrscheinlich ist – sehr hohe Anforderungen an die Ähnlichkeit zwischen Serien gestellt werden.

Bei kleineren Schwellen für die t-Werte kann laut Simulation im Untersuchungsgebiet nicht mehr garantiert werden, dass bestandesspezifische oder regionale Referenzchronologien aufgebaut werden. Jedoch scheint es bereits mit einem Schwellenwert von $t \geq 5$ und mit historischen Einzelserien möglich, eine Referenzchronologie für die Höhenlagen sowie eine für die Tief- und Mittellagen aufzubauen.

In Dendro-Provenancing Studien wird oft von regionalem Zuwachssignal gesprochen bzw. anhand weniger, stichprobenhafter Untersuchungen eine Intuition für "starke" oder "schwache" Signalunterschiede gebildet (Bernabei/Bontadi, 2011; Daly, 2007a; Eckstein/Wrobel, 2007; Grabner et al., 2018; Ważny, 2002). Es ist jedoch äusserst schwierig genau zu quantifizieren, ob und unter welchen Umständen regionale Zuwachssignale genügend spezifisch für das Dendro-Provenancing sind. Die Frage, wie hoch ein t-Wert sein muss, damit das in Kauf genommene Kontaminationsrisiko regionaler Referenzchronologien möglichst gering ist, steht im Zentrum einer klassischen Dendro-Provenancing Studie (Kap. 2.2). Für diese Problemstellung bietet der im Rahmen der vorliegenden Arbeit entwickelte Simulationsansatz bislang die einzige, methodische Herangehensweise, die es erlaubt Schwellenwerte quantitativ herzuleiten und dabei die zu erwartende Auflösung (bestandesspezifisch, lokal bis regional oder höhenlagenspezifisch) von Referenzchronologien im Untersuchungsgebiet auszuloten. Zudem liessen sich durch eine Modifikation des PREF-Constructor Algorithmus' künftig auch andere statistische Klassifikationsmethoden (als die 1-Nächste Nachbar Klassifikation) und weitere Ähnlichkeitsmasse (nebst der t-Wert Berechnung) implementieren.

6 Fallbeispiel: Holzherkunft der dendrochronologisch untersuchten, neuzeitlichen Gebäude der Stadt Zürich

6.1 Einführung

Die methodischen Entwicklungen des Dissertations-Projektes erlauben eine Neueinschätzung des Potentials für dendrochronologische Herkunftsbestimmungen historischer Zuwachsserien im Untersuchungsgebiet. Im selben Zug wird es möglich, die Ausgangslage für zukünftige Studien zu skizzieren.

Der Erfolg des Dendro-Provenancings in einem Untersuchungsgebiet ist primär von zwei Voraussetzungen abhängig: Erstens muss die Datengrundlage ausreichen, um

lokale Referenzchronologien aufzubauen. Zweitens müssen die Zuwachssignale allfällig etablierter Referenzchronologien so unterschiedlich sein, dass die Vermischung von Zuwachsserien verschiedener Herkunft unwahrscheinlich ist. Nur wenn beide dieser Bedingungen für ein Gebiet, wie hier das Einzugsgebiet der Holzversorgung der neuzeitlichen Stadt Zürich, erfüllt sind, kann die geographische Herkunft historischer Hölzer mittels Dendro-Provenancing unabhängig von historischen Schriftquellen bestimmt werden. Deshalb muss für das Untersuchungsgebiet und den Untersuchungszeitraum (die Jahre 1500 bis 1850) zunächst geklärt werden, welche Art von Referenzchronologien (Lokalchronologien, Regionalchronologien oder Höhenlagenchronologien) auf der derzeitigen Datengrundlage aufgebaut werden können.

Aufgrund der häufigen Verwendung von Fichte als historisches Bauholz sind zu Beginn des Dissertations-Projekts rezente Referenzchronologien für diese Baumart in der Nähe historischer Holzeinschlagsgebiete aufgebaut worden (vgl. Kap. 3).⁸ Die in diesem Rahmen etablierten, rezenten Fichtenchronologien sind die bislang einzigen, welche spezifisch für das Dendro-Provenancing im Untersuchungsgebiet erstellt worden sind. Deshalb fehlen entsprechende Referenzchronologien für andere Baumarten. Da bei der Erweiterung rezenter Chronologien mittels historischer Serien keine Vermischung der Baumarten erfolgen darf (Schweingruber, 1983, 131), beschränkt sich die folgende Untersuchung auf die Baumart Fichte. Die Darstellung der entsprechenden Datengrundlage bildet den Auftakt. Darauf folgen eine knappe Methodenbeschreibung sowie die Präsentation der Ergebnisse. Den Abschluss des Kapitels bildet die Diskussion der erstellten Referenzchronologien sowie eine Einschätzung des derzeitigen Potentials des Dendro-Provenancings im Untersuchungsgebiet.

6.2 Dendrochronologische Datengrundlage

Fichtenserien mit weniger als 50 Jahrringen sind zu kurz für den Aufbau von Referenzchronologien, weil die statistische Datenbasis für die Identifikation des

⁸ Im Zuge von Artikel II wurde eine weitere Chronologie im Bestand Albisboden aufgebaut (Kürzel ABPA in Artikel II). Jedoch waren alle bei dieser Aktion beprobten Fichten unter 50 Jahre alt. Es handelt sich somit um eine Chronologie, die für die Fragestellung dieses Kapitels zu wenige Jahrringe aufweist (vgl. Kap. 6.2). Deshalb werden die Daten aus dem Bestand Albisboden hier nicht integriert. Mit "rezenten Beständen" sind deshalb im Folgenden die im Zuge des Artikels I erstellten Bestandeschronologien gemeint.

lokalen Signals zu gering ist (Fowler/Bridge, 2017; Hollstein, 1980; Wigley et al., 1987). Den Ausführungen in Artikel III folgend wird deshalb eine Mindestüberlappung von 50 Jahren für alle Vergleiche zwischen Serien gefordert. Dies ist insofern bedeutend, als 30- bis 50-jährige Fichten im neuzeitlichen Baumaterial sehr häufig sind, jedoch nun aus dem Datensatz fallen.⁹ Folglich ist hervorzuheben, dass rein aufgrund der statistisch notwendigen Mindestanzahl Jahrringe immer nur eine Auswahl des Baumaterials überhaupt klassifizierbar ist.

Aufgrund der getroffenen Abgrenzung des Untersuchungsgebiets und -zeitraums (siehe Kap. 3) beschränkt sich die folgende Analyse auf Zuwachsserien mit dem Enddatum im 16. Jh. und später. Im Labor der Stadt Zürich finden sich – mit unterschiedlichen regionalen Schwerpunkten – dendrochronologische Daten aus der ganzen Schweiz und dem angrenzenden Ausland. Wie bereits ausgeführt, interessieren hier besonders die Fichtenserien aus dem Einzugsgebiet für die Holzversorgung der neuzeitlichen Stadt Zürich. Dieses umfasst Teile der Kantone Zürich, Schwyz, Glarus, St. Gallen, Zug und Uri (vgl. Kap. 3).

Der Datenbestand von eindeutig datierten historischen Fichtenzuwachsserien (≥ 50 Jahrringe) aus den oben genannten Kantonen und mit Endjahr im Untersuchungszeitraum umfasst 1702 Einzelserien (Abb. 1; graue Belegung).¹⁰ Ca. 73% dieser Serien (1235) stammen von Untersuchungen auf dem Gebiet der Stadt Zürich (Abb. 1; schwarze Belegung). Da besonders für die Hölzer aus Stadtschweizer Bauobjekten aufgrund möglicher Flossholznutzung die Holzentnahme aus demselben Bestand sehr unwahrscheinlich ist, sind keine Gebäudechronologien verfügbar, die für das Dendro-Provenancing geeignet wären (vgl. Artikel III).

9 Freundliche mündliche Mitteilung Felix Walder, UWAD ZH.

10 Für den groben Überblick werden hier alle Daten aus den genannten Kantonen, die den eben definierten Kriterien entsprechen und die im Labor der Stadt Zürich verfügbar waren (Stand 16.8.2019), beigezogen. Auf den Ausschluss der Daten, welche aus Regionen stammen, die nicht an das hydrologische System oberhalb der Stadt Zürich angeschlossen sind, wird an dieser Stelle verzichtet (Begründung weiter unten im Text).

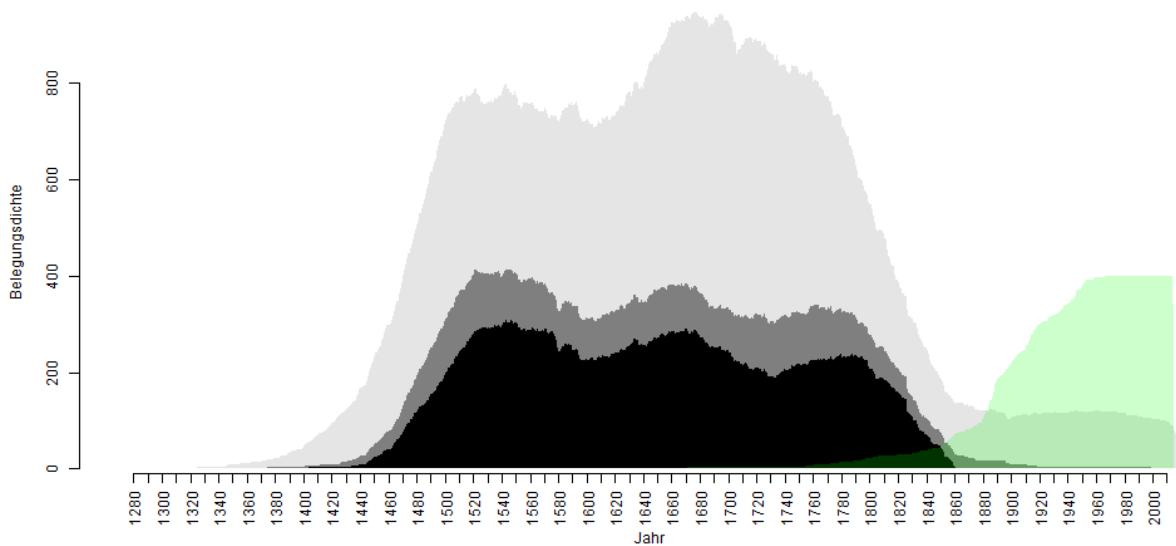


Abb. 1: Belegungsdichte der Zuwachsseries. Schwarz: Serien aus Untersuchungen in der Stadt Zürich; grau: Serien aus dem engeren Untersuchungsgebiet; hellgrau: Erweiterter Datensatz; grün: Rezente Serien aus Artikel I.

Die Belegungsdichte weist über den Untersuchungszeitraum hinweg eine inhomogene Verteilung auf: Während der Zeitraum 16. - 18. Jh. relativ gleichmässig belegt ist, ist die Belegung ab Mitte des 19. Jh. im Untersuchungsgebiet äusserst schwach (Abb. 1; schwarze Belegung). Der Grund dafür liegt in der Einführung der obligatorischen kantonalen Gebäudeversicherung um 1808 im Kanton Zürich (Körner/Degen, 2011). Deshalb sind ab Beginn des 19. Jh. Bauphasen zunehmend über Versicherungsakten erschliessbar. Dementsprechend selten werden von denkmalpflegerischer Seite Aufträge für dendrochronologische Untersuchungen in solchen Gebäuden vergeben. Selbst für die gut untersuchte Stadt Zürich ist eine zuverlässige Überbrückung zwischen rezenten Beständen und historischen Gebäuden deshalb nur möglich, wenn alle im Untersuchungsgebiet verfügbaren Zuwachsseries zusammengefasst betrachtet werden (Abb. 1; graue Belegung).

Die restlichen Regionen sind deutlich weniger umfassend untersucht. Die räumliche und zeitliche Inhomogenität des Datensatzes verhindert daher selbst nach Jahrzehnten intensiver dendrochronologischer Tätigkeit den separaten Aufbau von durchgehenden und ausreichend belegten Lokalchronologien in den einzelnen Regionen des Untersuchungsgebiets. Die bestehenden Gebäudechronologien wurden zudem für die Datierung und nicht fürs Dendro-Provenancing erstellt. Sie alle

müssten eingehend auf ihre beständliche Zusammensetzung und Eignung fürs Dendro-Provenancing überprüft und gegebenenfalls neu aufgebaut werden.¹¹ Dies zu leisten, übersteigt den verbleibenden Zeitrahmen des Dissertations-Projektes, weshalb lediglich Einzelserien für die folgende Analyse bereitstehen.

Bei der Verwendung von Einzelserien ist aber gemäss den Resultaten aus Artikel III auch mit einem sehr homogenen und grossen Datensatz kein Aufbau von Lokalchronologien möglich. Das Kontaminationsrisiko ist zu hoch (vgl. Resultate Einzelserie-Ansatz Kap. 5.3.2 sowie Artikel III). Folglich entstünden viele, potenziell falsche Provenancing-Resultate, die nicht räumlich zu interpretieren wären. Weniger problematisch ist vor dem Hintergrund der Simulationsergebnisse der Aufbau von Höhenlagenchronologien (vgl. Kap. 5.3.3). Besonders praktikabel dürfte der Aufbau einer Chronologie für die Hochlagen (>1500 m ü. M.) sowie der Aufbau einer gemeinsamen Chronologie für Tief- und Mittellagen (<1500 m ü. M.) sein. Deshalb wird hier nur der Aufbau entsprechender Höhenlagenchronologien bzw. eine Klassifikation in Hoch- und Tief-/Mittellagen weiterverfolgt.

Die geographische Verteilung der Daten aus dem Untersuchungsgebiet ist stark Richtung Stadt Zürich verzerrt (Abb. 1). Gleichzeitig ist unbekannt, aus welchen Höhenlagen die in Zürich verbauten Hölzer stammen. Um bessere und für alle Höhenlagen gleichmässigere Voraussetzungen für die Überbrückung von rezenten Chronologien und historischen Serien zu schaffen, werden alle im Zürcher Labor verfügbaren Fichtenserien in die hiesige Analyse integriert. Hinzu kommt ein Datensatz, der freundlicherweise von Matthias Bolliger (Archäologischer Dienst des Kantons Bern, Dendrolabor) zur Verfügung gestellt wurde. Dank dieser Erweiterung des Datensatzes stammen die für den Referenzchronologieaufbau verwendeten Serien nun aus dem gesamten Mittelland sowie den nördlichen Voralpen und Alpen der Schweiz. Hinzu kommen vier Serien aus Österreich; ebenfalls nördlicher Alpenraum. Serien südlich der Alpen und solche, die nicht aus dem alpinen bzw. zirkumalpinen Raum stammen, wurden aus dem Datensatz entfernt. Es wurden nur Serien berücksichtigt, welche die oben aufgeführten Kriterien (≥ 50 Ringe, sicher datiert sowie ein

¹¹ Die genannten 1702 Serien aus dem engeren Untersuchungsgebiet verteilen sich bereits auf 408 untersuchte Objekte bzw. unterschiedliche Bauphasen desselben Objektes. Es ist also mit mindestens ebenso vielen Mittelkurven zu rechnen, die es im Detail zu überprüfen gäbe.

Endjahr im Jahr 1500 oder später) erfüllen. Insgesamt enthält der erweiterte Datensatz 3818 Fichtenserien (Abb. 1; hellgraue Belegung).

6.3 Methoden

Da das Ziel die Aufgliederung von Höhenlagensignalen war, konnten die rezenten Bestandeschronologien zu zwei höhenlagspezifischen Initialchronologien zusammengefasst werden. Jedoch war es aufgrund der mitunter grossen Unterschiede in den Bestandessignalen, die in Artikel I festgestellt worden waren, denkbar, dass sich die Bestandessignale bei der Mittelung aller Hoch- bzw. aller Tief- und Mittellagenserien gegenseitig aufheben. Besonders in den Tief- und Mittellagen war unklar, ob ein genügend starkes, überregionales Signal für den Referenzaufbau existierte. Bei den Hochlagenserien hingegen wurde aufgrund der Simulationsergebnisse (Artikel III) bereits vermutet, dass ein gemeinsames Hochlagensignal bestand (siehe auch Ergebnisse früherer Untersuchungen in Hochlagen: Wilson/Hopfmüller, 2001; Dittmar et al., 2012; Savva et al., 2006).

Mittels einer für reale Zuwachsserien angepassten Version des in Artikel III entwickelten Algorithmus' (PREF-Constructor, vgl. Artikel III) wurden deshalb zwei unterschiedliche Ansätze für den Referenzchronologieaufbau verfolgt: Zunächst wurden lediglich zwei höhenlagspezifische Initialchronologien aus allen Hoch- bzw. Tief- und Mittellagenserien erstellt und dann versucht entsprechend zwei Referenzen aufzubauen. Danach wurden – wie in der Simulation – alle 15 rezenten Bestandeschronologien aus Artikel I als Initialchronologien verwendet. Die binäre Zuordnung zu Hoch- bzw. Tief- und Mittellagen wurde in diesem Fall erst nach Aufbau aller (bis zu 15) Referenzen vollzogen. Dazu wurden alle Serien, die einer mittels Tief- oder Mittellagen-Bestandeschronologie initialisierten Referenz zugeordnet worden waren, als Tief- und Mittellagenserien klassifiziert. Ebenso wurden alle Serien, die einer durch Hochlagen-Bestandeschronologien initialisierten Referenz zugewiesen worden waren, als Hochlagenserien klassifiziert.

In Artikel III besaßen die simulierten Serien statistische Eigenschaften, welche denen autoregressiver Residualserien sehr ähnlich sind. Deshalb, und weil in Artikel I bereits gezeigt worden ist, dass sich das für das Dendro-Provenancing relevante Signal durch diese Art der statistischen Vorbehandlung verstärken lässt, wurden für die historischen wie die rezenten Serien autoregressive Residualserien berechnet (Berechnung siehe Artikel I). Lediglich diese Residualserien wurden vom Algorithmus

weiterverwendet. Die Erweiterung der rezenten Initialchronologien mittels historischer Serien erfolgte gemäss der jeweils Nächsten Nachbarn (des einen Nächsten Nachbarn, siehe Artikel I und III). Wiederum galten für die paarweisen Vergleiche zwischen Initialchronologie und historischen Serien Mindestüberlappungen von 50 Jahren. Gemäss den Resultaten in Artikel III wurde der Schwellenwert auf $t \geq 5$ gesetzt, welcher dort ausreichend war für die Unterscheidung von Hochlagen- und Tief-/Mittellagenserien (vgl. Fig 2, Artikel III).

Um bei der ursprünglichen Fragestellung zu bleiben, wurde nur die Höhenlagenzuweisung der historischen Serien aus der Stadt Zürich genauer untersucht. In anderen Regionen des Untersuchungsgebietes erlaubt es der Arbeitsstand aufgearbeiteter historischer Quellen nicht, Deutungsversuche für allfällig sich abzeichnende Befunde anzustellen (Weisz et al., 1983). Ebenfalls wurden Serien mit Endjahr nach 1850 nicht analysiert, da der Bauholztransport ab dem Zeitpunkt auch per Bahn möglich gewesen wäre und somit Holz auch von ausserhalb des Untersuchungsgebiets (siehe Kap. 3) hätte eingeführt werden können (Dubler, 2011).¹²

Durch Hinzunahme der Daten aus dem Berner Labor gelangten 91 neue, rezente Serien mit bekannter Höhenlage in den Datensatz. Bei diesen konnte deshalb überprüft werden, ob Serien, die der Hoch- bzw. Tief- und Mittellagenchronologie zugewiesen wurden, auch tatsächlich aus entsprechenden Höhenlagen stammten.

6.4 Resultate

6.4.1 Ansatz I: Zwei höhenlagspezifische Initialchronologien

Bei der Verwendung von zwei Initialchronologien, einer für die Hoch- und einer für die Tief- und Mittellagen, wurden rund 10% der Serien im Gesamtdatensatz einer der beiden Referenzchronologien zugewiesen. Vom Stadtschreiber-Datensatz sind 3% der historischen Zuwachsserien (36 Serien) klassifiziert worden, alle als Hochlagenserien (Tab. 1). Die Belegungsdichte der entstandenen Hochlagenchronologie war über den Untersuchungszeitraum hinweg relativ homogen, wobei die Belegung zu Beginn und Ende des Zeitraums etwas schwächer war (Abb. 2, Abb. 3). Die aufgebaute Tief- und Mittellagenchronologie war sehr schwach belegt (7 Serien). Darunter befand sich keine Serie aus der Stadt Zürich.

¹² Aufgrund dieser Begrenzung verringerte sich die Anzahl der genauer betrachteten Serien von 1235 auf 1190.

Tab. 1: Überblick der Klassifikationen, die aus den zwei verfolgten Ansätzen beim Aufbau von Referenzchronologien für die neuzeitliche Stadt Zürich hervorgingen: “2_init.chronos” = Ansatz I; “15_init.chronos” = Ansatz II; “n_tot” = Anzahl Serien im Gesamtdatensatz; “n_Zü” = Anzahl Serien im Datensatz der Stadt Zürich mit Endjahr in den Jahren 1500 bis 1850; “nkl_tot” = Anzahl der klassifizierten Serien im Gesamtdatensatz; “nkl_Zü” = Anzahl klassifizierter Serien in der Stadt Zürich; “%kl_Zü” Prozentsatz der Stadtzürcher Serien, welche klassifiziert wurden; “nH_Zü” = Anzahl Hochlagenserien in Zürich; “nMT_Zü” = Anzahl Tief- und Mittellagenserien in Zürich; “Obj_Zü” = Anzahl der Stadtzürcher Objekte in denen Serien klassifiziert wurden; “ObjZü_H” = Anzahl der Stadtzürcher Objekte in denen Hochlagenserien gefunden wurden; “ObjZü_MT” = Anzahl der Stadtzürcher Objekte in denen Tief- und Mittellagenserien gefunden wurden; “ObjZü_be” = Anzahl der Stadtzürcher Objekte in denen Tief- und Mittellagenserien sowie Hochlagenserien gefunden wurden.

	n_tot	n_Zü	nkl_tot	nkl_Zü	%kl_Zü	nH_Zü	nMT_Zü	Obj_Zü	ObjZü_H	ObjZü_MT	ObjZü_be
2_init.chronos	3818	1190	383	36	3	36	0	24	24	0	0
15_init.chronos	3818	1190	1386	363	31	36	327	144	17	120	7

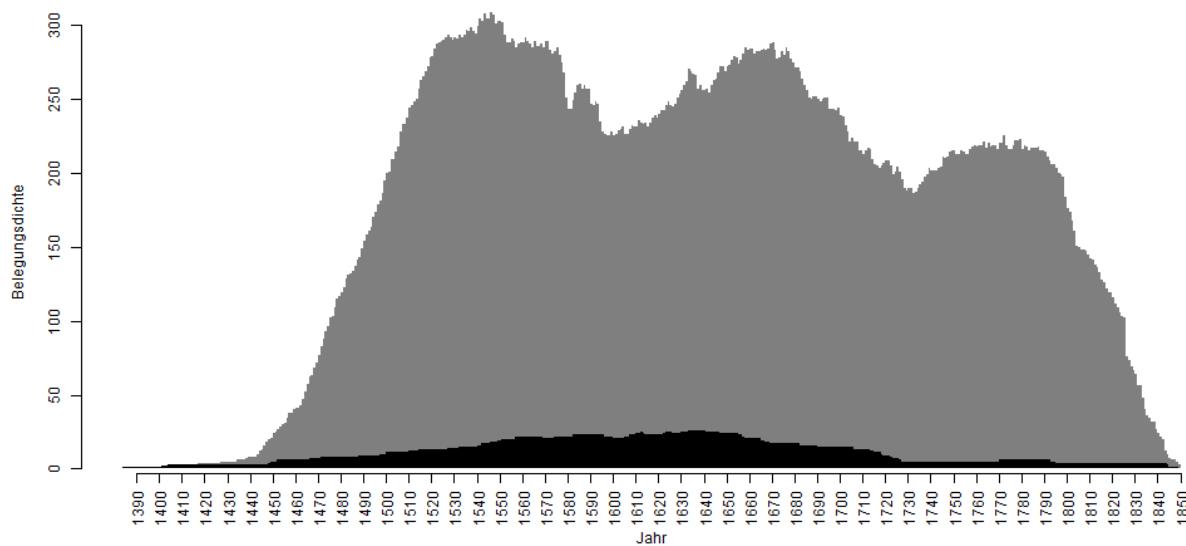


Abb. 2: Belegungsdichte der Zuwachsserien. Schwarz: Hochlagenserien aus Untersuchungen in der Stadt Zürich; grau: Nichtklassifizierte Serien aus dem Datensatz der Stadt Zürich.

Die der Hochlagenchronologie zugewiesenen, historischen Serien aus der Stadt Zürich verteilten sich auf insgesamt 24 Objekte. Ein Drittel davon (8 Objekte) waren Musikinstrumente. Der Rest, also 16 Objekte, stammte aus historischen Gebäuden (Tab. 1). Jedoch existieren für diese Gebäude derzeit keine aufgearbeiteten, historischen Quellen, die weitere Aussagen zu dem beim Bau verwendeten Holz erlaubten (vgl. Weisz et al., 1983; Bitterli, 2009).

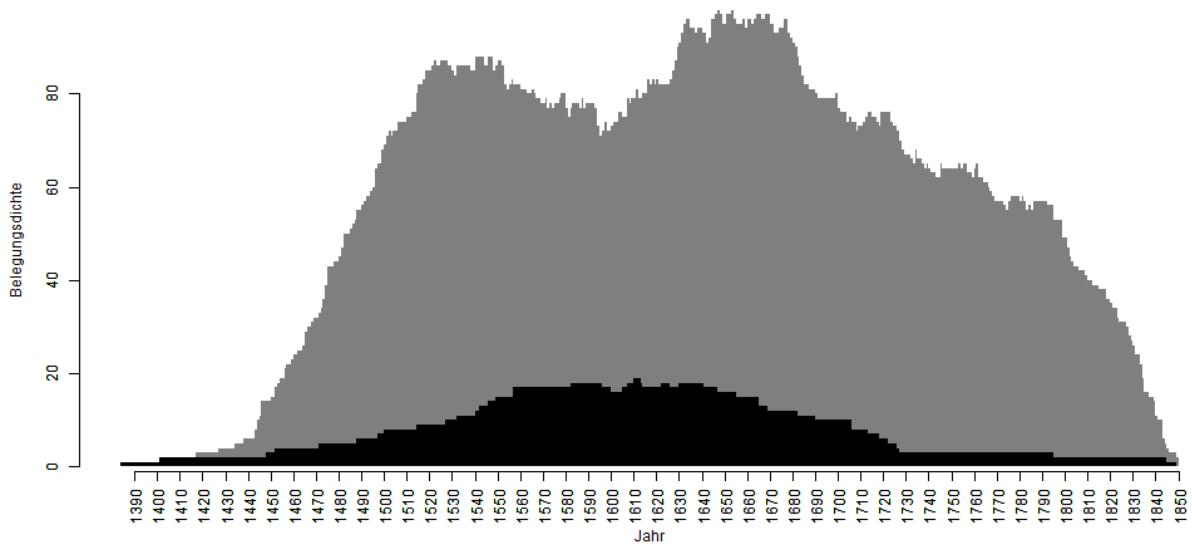


Abb. 3: Belegungsdichte der Zuwachsserien bereinigt nach Objekten. Schwarz: Belegungsdichte der Objekte mit Hochlagenserien aus Untersuchungen in der Stadt Zürich; grau: Belegungsdichte der Objekte mit nichtklassifizierten Serien aus dem Datensatz der Stadt Zürich.

Von den neuen, rezenten Serien im Datensatz, die nicht bereits in die Initialchronologien einflossen, sind nur Serien mit einer Herkunft aus 1300 m ü. M. und höher der Hochlagenchronologie zugewiesen worden. In der Tief- und Mittellagenchronologie befanden sich ausschliesslich Serien aus Beständen unterhalb von 1000 m ü. M.

6.4.2 Ansatz II: 15 Bestandeschronologien als Initialchronologien

Wurden alle rezenten Bestandeschronologien zur Initialisierung verwendet, stieg die Klassifikationsrate gegenüber Ansatz I stark an. Im Gesamtdatensatz wurden 36% der Serien klassifiziert. Von den Serien aus der Stadt Zürich wurden 31% einer Referenzchronologie zugewiesen (Tab. 1). Die Referenzen, die mit den Beständen *furg* bzw. *nb* (Bestandeskürzel siehe Artikel I, Tab. 1) initialisiert worden waren, wiesen die weitaus höchste Belegung auf (Tab. 2). Interessanterweise zeigte sich nach Aufsummierung aller Serien, dass solche, die einer der drei mittels Hochlagenchronologien initialisierten Referenzen (*furg*, *rw*, *ww*) zugewiesen worden waren, dieselben waren, die auch mit Ansatz I als Hochlagenserien klassifiziert worden waren. Im Gegensatz zum ersten Ansatz wurde jedoch eine grosse Anzahl Serien Referenzen zugewiesen, die mittels Tief- oder Mittellagenbestandeschronologien initialisiert worden waren (Tab. 2). Aufsummiert waren folglich 327 Serien als Tief- und Mittellagenserien klassifiziert worden.

Tab. 2: Anzahl Serien pro Referenzchronologie, die entsprechend mittels rezenter Bestandeschronologie initialisiert worden sind. Die Kürzel der Bestandeschronologien sind in Artikel I Tab. 1 aufgelöst. “total” = Klassifikationen im Gesamtdatensatz. “Zürich” = Klassifikationen im Datensatz der Stadt Zürich.

	sw	hw	furg	gw	rw	gand	ww	chw	kar	fri	how	nb	bw	sb	ew
total	1	0	328	2	24	2	37	3	0	0	1	982	3	3	0
Zürich	0	0	35	0	0	0	1	0	0	0	0	327	0	0	0

Während die Belegungsdichte der Hochlagenserien gleich blieb, besaß die Belegungsdichte der Tief- und Mittellagenserien zwei Maxima (ca. in den Jahren 1530 und 1680). Dazwischen, etwa in der Zeitspanne, in welcher die Hochlagenchronologie am dichtesten belegt war (1560 bis 1650), sank die Belegungsdichte der Tief- und Mittellagenserien zwischenzeitlich ab (Abb. 4). Diese Beobachtung wurde noch deutlicher, nachdem die Belegungsdichten nach Objekten bereinigt worden waren (Abb. 5).

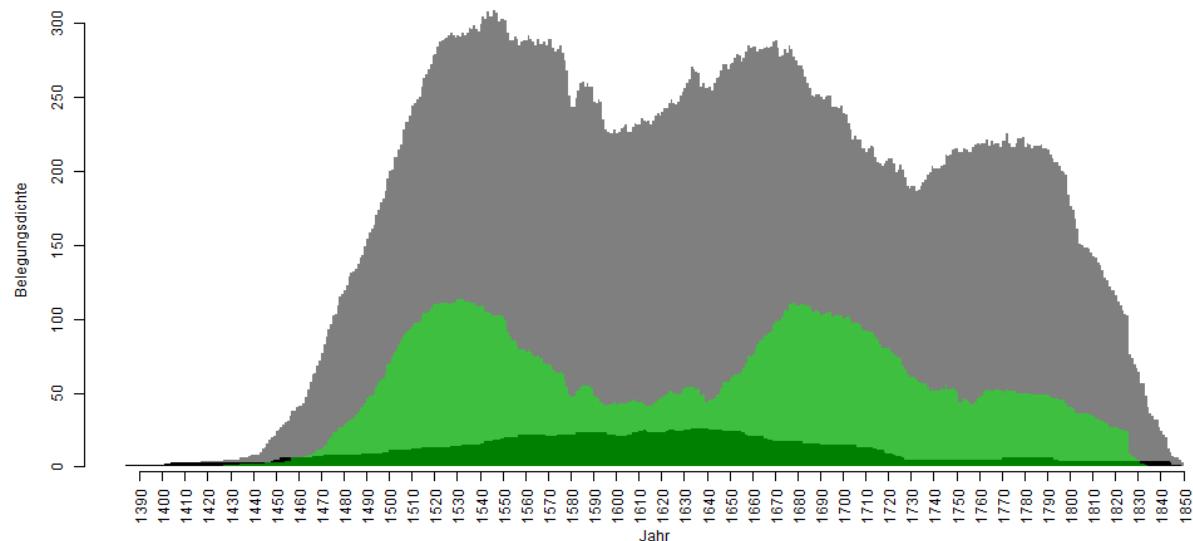


Abb. 4: Belegungsdichte der Zuwachsserien. Schwarz: Hochlagenserien aus Untersuchungen in der Stadt Zürich; Grün: Tief- und Mittellagenserien aus Untersuchungen in der Stadt Zürich; Grau: Nichtklassifizierte Serien aus dem Datensatz der Stadt Zürich.

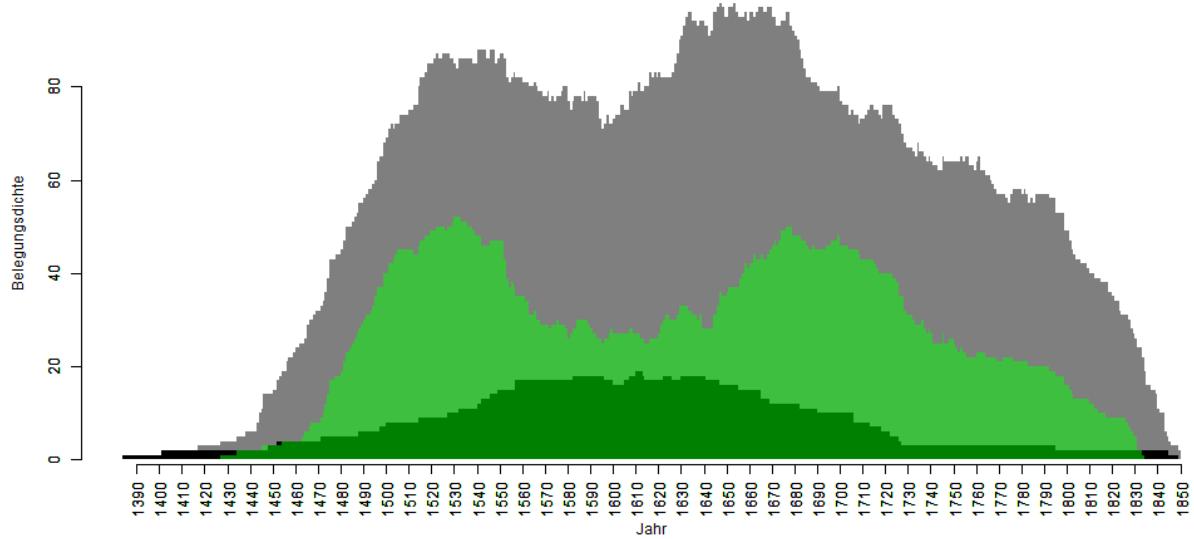


Abb. 5: Belegungsdichte der Zuwachsserien bereinigt nach Objekten. Schwarz: Belegungsdichte der Objekte mit Hochlagenserien aus Untersuchungen in der Stadt Zürich; Grün: Belegungsdichte der Objekte mit Tief- und Mittellagenserien aus Untersuchungen in der Stadt Zürich; Grau: Belegungsdichte der Objekte mit nichtklassifizierten Serien aus dem Datensatz der Stadt Zürich.

Die den Tief- und Mittellagen zugewiesenen Serien verteilten sich auf 120 Objekte, bei welchen die klassifizierten Serien nur dieser Höhenlagenklasse angehörten, sowie 7 Objekte, in denen sich zusätzlich auch Hochlagenserien befanden. Die Klassifikation der Hochlagenserien änderte sich mit dem zweiten Ansatz nicht. Abzüglich der 7 Objekte, in denen beide Höhenlagenklassen vorkamen, blieben folglich 17 Objekte, in welchen lediglich Hochlagenserien vorgefunden wurden. Darunter wiederum dieselben 8 Musikinstrumente, die bereits oben erwähnt worden sind (Tab. 1, Tab. 3).

Tab. 3: Vollständige Tabelle der Objekte, in denen gemäss Klassifikation mit Ansatz II Hoch- bzw. Tief- und Mittellagenserien vorgefunden wurden.

Objekt	Klasse der Serien
ZH/ZUERICH-GEIGE	Hoch
ZH/ZUERICH-En/KAPPELISTRASSE 11	Hoch
ZH/ZUERICH-FLORHOFGASSE 11	Hoch
ZH/ZUERICH-GEIGE	Hoch
ZH/ZUERICH-GEIGE	Hoch
ZH/ZUERICH-GEIGE	Hoch
ZH/ZUERICH-Hi/HAMMERSTRASSE 54	Hoch
ZH/ZUERICH-MUENSTERHOF 18	Hoch
ZH/ZUERICH-NIEDERDORFSTRASSE 86	Hoch
ZH/ZUERICH-BAROCKLAUTE	Hoch
ZH/ZUERICH-THEORBE	Hoch
ZH/ZUERICH-RENNWEG 7	Hoch

ZH/ZUERICH-RINDERMARKT 11	Hoch
ZH/ZUERICH-RINDERMARKT 7	Hoch
ZH/ZUERICH-SCHMIDGASSE 4	Hoch
ZH/ZUERICH-GEIGE	Hoch
ZH/ZUERICH-VIOLA	Hoch
ZH/ZUERICH-Af/KORNAMTSWEG 2	MittelTief
ZH/ZUERICH-Af/KORNAMTSWEG BEI 8	MittelTief
ZH/ZUERICH-Ar/ALBISRIEDERSTR.410	MittelTief
ZH/ZUERICH-Ar/ALBISRIEDERSTRASSE 362	MittelTief
ZH/ZUERICH-Ar/ALBISRIEDERSTRASSE 400	MittelTief
ZH/ZUERICH-Ar/ALBISRIEDERSTRASSE 401	MittelTief
ZH/ZUERICH-Ar/IM KRATZ 5	MittelTief
ZH/ZUERICH-Ar/LYRENWEG 8	MittelTief
ZH/ZUERICH-As/KANZLEISCHULHAUS/VORDACH	MittelTief
ZH/ZUERICH-AUGUSTINERGASSE 28	MittelTief
ZH/ZUERICH-AUGUSTINERGASSE 38	MittelTief
ZH/ZUERICH-AUGUSTINERGASSE 46	MittelTief
ZH/ZUERICH-CHORGASSE 15	MittelTief
ZH/ZUERICH-En/BEDERSTRASSE 65	MittelTief
ZH/ZUERICH-En/BRANDSCHENKESTRASSE 110-118	MittelTief
ZH/ZUERICH-En/BRANDSCHENKESTRASSE 114	MittelTief
ZH/ZUERICH-En/BRANDSCHENKESTRASSE 114/118	MittelTief
ZH/ZUERICH-En/BRANDSCHENKESTRASSE 118	MittelTief
ZH/ZUERICH-En/BRANDSCHENKESTRASSE 146	MittelTief
ZH/ZUERICH-En/GABELRSTRASSE 35/37	MittelTief
ZH/ZUERICH-Fi/FREUDENBERGSTRASSE 92	MittelTief
ZH/ZUERICH-FRANKENGASSE 11	MittelTief
ZH/ZUERICH-Hg/ACKERSTEINSTRASSSE 207/209/LIMMATTALSTRASSE 181	MittelTief
ZH/ZUERICH-Hg/AM HOENGGERBERG 110	MittelTief
ZH/ZUERICH-Hg/AM WASSER 48	MittelTief
ZH/ZUERICH-Hg/GAESSLI 3	MittelTief
ZH/ZUERICH-Hg/WETTINGERTOBEL 38/PFARRHAUS	MittelTief
ZH/ZUERICH-Hg/WINZERHALDE 51	MittelTief
ZH/ZUERICH-Hi/BURENWEG 12	MittelTief
ZH/ZUERICH-Hi/BURENWEG 22/24	MittelTief
ZH/ZUERICH-Hi/EIERBRECHTSTRASSE 42-46	MittelTief
ZH/ZUERICH-Hi/FORCHSTRASSE 246	MittelTief
ZH/ZUERICH-Hi/FORCHSTRASSE 248	MittelTief
ZH/ZUERICH-Hi/HAMMERSTRASSE BEI NR.84	MittelTief
ZH/ZUERICH-Hi/HEGIBACHSTR.84	MittelTief
ZH/ZUERICH-Hi/HEGIBACHSTRASSE 130/132/134	MittelTief
ZH/ZUERICH-Hi/WASERSTRASSE 18	MittelTief
ZH/ZUERICH-Hi/WASERSTRASSE 36	MittelTief
ZH/ZUERICH-Ho/ADLISBERGSTRASSE 70/72	MittelTief
ZH/ZUERICH-Ho/HOFSTRASSE 108	MittelTief
ZH/ZUERICH-Ho/HOFSTRASSE 31	MittelTief
ZH/ZUERICH-Ho/HOFSTRASSE 33	MittelTief
ZH/ZUERICH-Ho/HOFSTRASSE 40	MittelTief
ZH/ZUERICH-Ho/HOFSTRASSE 7/9/11	MittelTief
ZH/ZUERICH-Ho/KREUZPLATZ 10/11	MittelTief
ZH/ZUERICH-Ho/KREUZPLATZ 6	MittelTief

ZH/ZUERICH-Ho/KREUZPLATZ 7	MittelTief
ZH/ZUERICH-Ho/TOBELHOFSTRASSE 234,236/TOBELHOF	MittelTief
ZH/ZUERICH-LIMMATQUAI 104	MittelTief
ZH/ZUERICH-LIMMATQUAI 18	MittelTief
ZH/ZUERICH-LINDENHOFSTRASSE 3	MittelTief
ZH/ZUERICH-MARKTGASSE 15	MittelTief
ZH/ZUERICH-MARKTGASSE 17	MittelTief
ZH/ZUERICH-MARKTGASSE 20/ZUNFTHAUS ZUR SCHMIDEN	MittelTief
ZH/ZUERICH-MUENSTERGASSE 14	MittelTief
ZH/ZUERICH-NAPFGASSE 4	MittelTief
ZH/ZUERICH-NEUMARKT 25	MittelTief
ZH/ZUERICH-NEUSTADTGASSE 11	MittelTief
ZH/ZUERICH-OETENBACHGASSE 13	MittelTief
ZH/ZUERICH-OETENBACHGASSE 5	MittelTief
ZH/ZUERICH-OETENBACHGASSE 7	MittelTief
ZH/ZUERICH-Os/CULMANNSTRASSE 27/29	MittelTief
ZH/ZUERICH-Os/UNIVERSITAETSTRASSE 111	MittelTief
ZH/ZUERICH-Os/UNIVERSITAETSTRASSE 91/93	MittelTief
ZH/ZUERICH-PREDIGERGASSE 8	MittelTief
ZH/ZUERICH-PREDIGERKIRCHE	MittelTief
ZH/ZUERICH-PREDIGERPLATZ 38	MittelTief
ZH/ZUERICH-Rb/DRAHTZUGSTRASSE 72	MittelTief
ZH/ZUERICH-Rb/FLUEHGASSE 12	MittelTief
ZH/ZUERICH-Rb/FLUEHGASSE 14	MittelTief
ZH/ZUERICH-Rb/FLUEHGASSE 27/29	MittelTief
ZH/ZUERICH-Rb/FLUEHGASSE/TROTTE BEI SEEFELDSTRASSE 277	MittelTief
ZH/ZUERICH-Rb/FORCHSTRASSE 303	MittelTief
ZH/ZUERICH-Rb/LENGGSTRASSE 7	MittelTief
ZH/ZUERICH-Rb/LUREIWEG 19	MittelTief
ZH/ZUERICH-Rb/SUEDSTRASSE 79/81	MittelTief
ZH/ZUERICH-RENNWEG 15	MittelTief
ZH/ZUERICH-RENNWEG 3	MittelTief
ZH/ZUERICH-Sb/AUSSERDORFSTRASSE 14	MittelTief
ZH/ZUERICH-Sb/AUSSERDORFSTRASSE 18/20	MittelTief
ZH/ZUERICH-Sb/BUHNRAIN 29-37	MittelTief
ZH/ZUERICH-Sb/BUHNSTRASSE 23	MittelTief
ZH/ZUERICH-SCHIFFLAENDE 14/16/SCHEITERGASSE 1	MittelTief
ZH/ZUERICH-SCHLUESSELGASSE 14	MittelTief
ZH/ZUERICH-SCHOFFELGASSE 11	MittelTief
ZH/ZUERICH-SPIEGELGASSE 18	MittelTief
ZH/ZUERICH-SPIEGELGASSE 2/NAPFGASSE 3	MittelTief
ZH/ZUERICH-SPIEGELGASSE 9	MittelTief
ZH/ZUERICH-ST.PETER	MittelTief
ZH/ZUERICH-STORCHENGASSE 12	MittelTief
ZH/ZUERICH-STREHLGASSE 17	MittelTief
ZH/ZUERICH-STUESSIHOFSTATT 4	MittelTief
ZH/ZUERICH-Sw/HEINRICH BOSSHARD-STRASSE 2 /WINTERTHURERSTRASSE 508	MittelTief
ZH/ZUERICH-Sw/HUETTENKOPFSTRASSE 11	MittelTief
ZH/ZUERICH-Sw/HUETTENKOPFSTRASSE 70/ZIEGELHUETTE	MittelTief
ZH/ZUERICH-Sw/STETTBACHSTRASSE 51	MittelTief
ZH/ZUERICH-Sw/WINTERTHURERSTRASSE 489	MittelTief

ZH/ZUERICH-Sw/WINTERTHURERSTRASSE 519	MittelTief
ZH/ZUERICH-TRITTLIGASSE 12	MittelTief
ZH/ZUERICH-TRITTLIGASSE 9	MittelTief
ZH/ZUERICH-U _s /ROESLISTRASSE 4	MittelTief
ZH/ZUERICH-Wd/FRIESENBERGSTRASSE 374-380	MittelTief
ZH/ZUERICH-Wd/HALDENSTRASSE 144	MittelTief
ZH/ZUERICH-Wd/HALDENSTRASSE 145	MittelTief
ZH/ZUERICH-WEGGENGASSE 3/SCHLUESSELGASSE 20	MittelTief
ZH/ZUERICH-WEINPLATZ 9-10/SCHIPFE 1	MittelTief
ZH/ZUERICH-Wo/ALBISSTRASSE 80	MittelTief
ZH/ZUERICH-Wo/ALTE KALCHBUEHLSTRASSE 21/23/TROTTGEBAEUDE	MittelTief
ZH/ZUERICH-Wo/ALTE KALCHBUEHLSTRASSE 21/23/WOHNHAUS	MittelTief
ZH/ZUERICH-Wo/GRETNENWEG 3	MittelTief
ZH/ZUERICH-Wo/SEESTRASSE 520	MittelTief
ZH/ZUERICH-Wp/NORDSTRASSE 331	MittelTief
ZH/ZUERICH-Wp/ROESCHIBACHSTRASSE 4/6/8/10	MittelTief
ZH/ZUERICH-W/BERGHALDENSTRASSE 71	MittelTief
ZH/ZUERICH-Wt/KAPFSTEIG 58	MittelTief
ZH/ZUERICH-Wt/MOECKLISTRASSE 1	MittelTief
ZH/ZUERICH-Wt/WITIKONERSTRASSE 394/MEIERHOF	MittelTief
ZH/ZUERICH-Wt/WITIKONERSTRASSE 421	MittelTief
ZH/ZUERICH-WUEHRE 13	MittelTief
ZH/ZUERICH-ZINNENGASSE 5	MittelTief
ZH/ZUERICH-Ar/ALTSTETTERSTRASSE 336	Beides
ZH/ZUERICH-BRUNNGASSE 8	Beides
ZH/ZUERICH-Hg/GSTEIGSTRASSE 9/11	Beides
ZH/ZUERICH-Hg/WIESLERGASSE 22	Beides
ZH/ZUERICH-OBERDORFSTRASSE 5	Beides
ZH/ZUERICH-Rb/FLUEHGASSE 9/11	Beides
ZH/ZUERICH-RENNWEG 2	Beides

Auch in den Referenzen, die mittels 15 Bestandeschronologien initialisiert worden waren, fanden sich Serien, die von rezenten Fichten von oberhalb 1300 m ü. M. stammten (Berner Material), praktisch ausschliesslich in entsprechenden Hochlagenreferenzen. Nur eine Serie wurde als Tief- und Mittellagenserie klassifiziert, obwohl der betreffende Baum eigentlich auf 1510 m ü. M. gewachsen war. Die aus den Tief- und Mittellagen (477 m bis 1009 m ü. M.) stammenden und klassifizierten, rezenten Berner Serien wurden alle Referenzen zugewiesen, die durch Tief- bzw. Mittellagenbestandeschronologien initialisiert worden waren.

6.5 Diskussion und Schlussfolgerungen

Mit dem Ansatz zweier Initialchronologien, einer für die Hoch- und einer für die Tief- und Mittellagenbestände, waren lediglich 3% der Serien aus der Stadt Zürich einer Referenz, nämlich der Hochlagenchronologie, zuweisbar. Die Tief- und Mittellagenchronologie liess sich mit Ansatz I *de facto* nicht aufbauen. Mit den 15 Bestandes-

chronologien hingegen gelang es insbesondere für die mittels der Bestandeschronologie von *nb* initialisierte Referenz eine lange Chronologie aufzubauen, die im Datensatz der Stadt Zürich 327 Serien umfasste. Bei den Hochlagenserien änderte sich für das städtische Material nichts. Das gemeinsame Signal zwischen den Beständen in den tieferen Lagen scheint somit weniger ausgeprägt zu sein als in den Hochlagen. Folglich haben sich die rezenten Bestandessignale vermutlich bei der Bildung einer gemeinsamen Initialchronologie für die Tief- und Mittellagenchronologie gegenseitig ausgemittelt.

Im Ansatz mit 15 Initialchronologien wurde diese “Ausmittelung” umgangen und es gelang prompt der Aufbau einer durchgehenden Chronologie für die mit *nb* initialisierte Referenz. Gemäss der Simulationsergebnisse (Artikel III) sollte diese Referenz jedoch nicht als Lokalchronologie für den Bestand *nb* gedeutet werden, da bei der Verwendung von Einzelserien und einem Schwellenwert von $t \geq 5$ ein grosses Kontaminationsrisiko für das lokale Bestandessignal besteht. Es darf aber aufgrund der in der Simulation festgestellten Separierung von Tief-/Mittel- und Hochlagenserien sowie der Analyse der Zuweisung der rezenten Serien aus dem Berner Datensatz davon ausgegangen werden, dass es sich bei der mit *nb* initialisierten Referenz um eine Tief- und Mittellagenreferenz handelt (Interpretation weiter unten).

Ebenfalls gilt es zu bedenken, dass mit den hier aufgebauten Bestandeschronologien nur ein Teil der im Untersuchungsgebiet vorkommenden Bestandessignale erfasst worden ist. Würden die Bestandeschronologien tatsächlich alle relevanten Bestandessignale abdecken, würde sich die Klassifikationsrate vermutlich derjenigen der Simulation annähern (zwischen 43% bis 61%, siehe Artikel III, Table 1).

Mit beiden Ansätzen gelangten in die aufgebaute Hochlagenchronologien auch Serien aus rezenten Beständen von unterhalb 1500 m ü. M. Da jedoch keine Serien von unterhalb 1300 m ü. M. einflossen, scheint sich dennoch eindeutig ein Hochlagensignal abzuzeichnen. Die Schwelle ≥ 1500 m ü. M. für das Hochlagenwuchs-signal ergab sich aus den in Artikel I untersuchten Daten. Dort fehlen jedoch Bestände zwischen 1200 m und 1500 m ü. M. Im neuen bernerischen Material sind diese Höhenstufen nun vertreten. Deshalb überrascht es nicht, dass sich die Schwelle von ≥ 1500 m ü. M. als Artefakt entpuppte.

Es muss folglich mit regionalen Schwankungen der absoluten Höhe über dem Meeresspiegel gerechnet werden, auf der sich die Trennung zwischen Hoch- und

Tiefmittellagensignal vollzieht. Dieser Befund stimmt überein mit den bereits seit langem festgestellten, regionalen Unterschieden für das potentiell natürliche Vorkommen der Fichte (Ellenberg, 1996, 28, 123, 334–348; Pott, 1993, 52–68). Das Wachstum der Fichte ist abhängig von Umweltbedingungen, die auf einer bestimmten Höhe über dem Meeresspiegel herrschen. Falls die Bedingungen in verglichenen geographischen Gebieten gleich, die Höhe über dem Meeresspiegel aber unterschiedlich ist, kann das Wuchssignal dennoch ähnlich ausfallen. Dies zeigt sich in der Zuordnung der rezenten Serien aus dem Berner Datensatz, die aus 1300 m ü. M. stammen und trotzdem ein ähnliches Wuchssignal aufweisen, wie die Fichten von oberhalb 1500 m ü. M. aus dem Datensatz der in Artikel I publizierten Bestände.

Der kleine Anteil an Serien, die in beiden Ansätzen der Hochlagenchronologie zugewiesen wurden, sowie die relativ hohe Zahl der mit Ansatz II den Tief- und Mittellagen zugewiesenen Serien könnte darauf hinweisen, dass in der neuzeitlichen Stadt Zürich selten Holz aus Lagen oberhalb von 1300 m ü. M. verbaut worden ist. Fichten wuchsen auch vor der Förderung durch den Menschen bereits in wechselndem Ausmass beigemischt an Standorten auf 900 bis 1000 m ü. M. und kamen oberhalb davon in der Neuzeit häufig vor (Weisz et al., 1983, 407–408). Spätestens ab Ende 18. Jh. sind für den tiefgelegenen Sihlwald einsetzende Fördermassnahmen für die Fichte belegt (Weisz et al., 1983, 407–408). Gleichzeitig war der Transport von Bauholz, das in Hochlagen geschlagen wurde, bis ins 19. Jh. hinein sehr aufwändig (Bürgi/Schuler, 2001). Generell scheint vom Anfang des 16. bis zum Ende des 18. Jh. der Hauptteil des für öffentliche Bauten in Zürich verwendeten Holzes über den Zürich- und Walensee verhandelt worden zu sein, wo auch – besonders um den Walensee – leichter zugängliche Fichtenbestände unterhalb von 1300 m ü. M. existierten (Weisz et al., 1983, 356–358).

Das schwache Maximum in der Belegungsdichte der Hochlagenserien zur Zeit des zwischenzeitlichen Minimums bei den Tieflagenserien könnte möglicherweise mit dem intensivierten Bezug von Glarner Holz ab Ende 16. bis Anfang 18. Jh. zusammenhängen (Weisz et al., 1983, 358–360). Der Befund ist jedoch sehr unsicher, da sich das Minimum der Tief- und Mittellagenserien bei Erweiterung der zur Initialisierung verwendeten rezenten Bestandeschronologien und der dadurch potentiell höheren Klassifikationsrate wieder auflösen könnte. Zudem hätte Zürich

auch aus anderen Regionen als dem Glarnerland Hochlagenholz beziehen können (Weisz et al., 1983, 351–360).

Dass es sich bei den als Hochlagenserien klassifizierten Serien tatsächlich um Hochlagenholz handeln könnte, legt der hohe Anteil an Serien nahe, die von Musikinstrumenten stammt. Für letztere wurde in historischer Zeit oft Holz aus Hochlagenbeständen verwendet (Bernabei/Bontadi, 2011).

Der weitere Aufbau der Höhenlagenchronologie könnte sich in der Stadt Zürich schwierig gestalten, sollte sich die Anzahl der überhaupt aus diesen Lagen (>1300 m ü. M.) verbauten Hölzer – wie oben ausgeführt – tatsächlich als begrenzt erweisen. Für die Tief- und Mittellagen wurde, über die Erweiterung der Chronologie von *nb*, eine recht gut belegte Referenz erstellt. Diese ist aber nur mit Vorbehalt als Referenz für das überregionale Tief- und Mittellagen Wuchssignal zu bezeichnen. Es dürfte sich eher um ein Signal handeln, das einige Bestände auf dieser Höhenlage teilen, andere jedoch nicht, andernfalls wäre die Verstärkung durch Mittelung aller Bestandeschronologien der Tief- und Mittellagen – wie mit Ansatz I versucht – erfolgreich gewesen. Möglicherweise repräsentiert die mittels *nb* aufgebaute Referenz ein gemeinsames Wuchssignal, dessen Ausprägung durch ähnliche Standortfaktoren (zusätzlich zur gleichen Höhenlage) bedingt wird. Dieser Frage müsste jedoch in einer weiteren Studie nachgegangen werden.

Es ist festzuhalten, dass es trotz des grossen Datensatzes derzeit nicht möglich ist, Lokalchronologien anhand historischer Fichtenserien aufzubauen. Die Datengrundlage ist in den Herkunftsgebieten zu lückenhaft. In der Simulation wurden zwischen 6312 und 6518 Pseudo-Objektchronologien erzeugt, die jeweils durch 6 Pseudo-Serien belegt waren (Artikel III, Table 1). Für die 15 Pseudo-Bestandessignale waren in der Simulation folglich zwischen 2524 und 2607 Serien erzeugt worden, die absolut gleichmässig über den Untersuchungszeitraum hinweg die 15 Pseudo-Bestandessignale replizierten. Von dieser Datenmenge ist man heute in Realität weit entfernt. Zudem ist die Anzahl der tatsächlich im Untersuchungsgebiet vorkommender Bestandessignale unbekannt. Die Zahl der Serien, die für den Aufbau von Lokalchronologien notwendig wären, könnte sich deshalb leicht vervielfachen.

7 Schlussbetrachtung

7.1 Synthese der methodischen Fortschritte und Resultate

Im weitesten Sinn entspricht Dendro-Provenancing der methodischen Suche nach den besten Übereinstimmungen zwischen dendrochronologischen Serien unbekannter Herkunft und Referenzserien bekannter Herkunft. Dieser Kern bleibt auch nach den Beiträgen des Dissertations-Projektes erhalten. Bisher ging die Konzeption der Dendro-Provenancing Methode jedoch nicht von expliziten Grundannahmen aus (vgl. Kap. 2.2 und 4). Mit der hier präsentierten Arbeit (Artikel I) liegt nun eine Formulierung der Annahmen vor. Demzufolge ist der Erfolg von Dendro-Provenancing (1.) davon abhängig, dass das Baumwachstum in einem Untersuchungsgebiet ausreichend variiert, es also zur Ausprägung regionaler oder lokaler Jahrringbreitenmuster kommt. Damit solche Muster aber identifiziert und Referenzchronologien gebildet werden können, muss angenommen werden, dass (2.) diese Ähnlichkeit über statistische Vergleichsrechnungen quantifizierbar ist und (3.) die grösste Ähnlichkeit zwischen Zuwachsserien in nächster geographischer Nachbarschaft festgestellt wird.

Besonders die erste Grundannahme kann nicht *a priori* für ein Untersuchungsgebiet gesetzt werden. Damit Dendro-Provenancing Resultate überhaupt plausibel interpretiert werden können, muss diese Annahme zuerst anhand rezenter Zuwachsserien evaluiert werden. Mit der Einführung von k-NN Klassifikationen, die mittels *scissor plots* visualisiert werden, wurde ein entsprechendes Evaluationsverfahren präsentiert (Artikel I). Zudem können mit demselben Verfahren die zweite und dritte Annahme in einem Gebiet überprüft werden, indem die Performanz verschiedener Ähnlichkeitsmasse verglichen wird und so die optimale statistische Vorbehandlungsmethode und Vergleichsrechnung bestimmt werden können. Auf diese Weise gelang es, das hochfrequente Zuwachssignal in einem transparenten, quantitativen Verfahren als ausschlaggebend für den Erfolg des Dendro-Provenancings herauszustellen. Zudem konnte kein Einwand gegen die Verwendung von Einzelradien beim Aufbau von Objektchronologien gefunden werden (Artikel II). Gemäss den Ergebnissen von Artikel II ist es unwahrscheinlich, dass das Dendro-Provenancing relevante hochfrequente Signal durch den Probenentnahmewinkel verzerrt wird und sich diese Verzerrung in den auf Basis der Objektchronologien aufgebauten Referenzchronologien fortsetzt.

Das Vorhandensein von bestandesspezifischen hochfrequenten Zuwachssignalen allein ist jedoch nicht ausreichend für erfolgreiche Herkunftsbestimmungen. Während es in der Tat einige Untersuchungen gegeben hat, die rezente Zuwachsserien auf Ähnlichkeitscluster hin untersucht haben, wurde in der Forschung bislang kaum thematisiert, dass es nicht zwingend zutrifft, dass sich regionale Zuwachsmuster beim Aufbau von Referenzchronologien erhalten (vgl. Kap. 5.3). In einem Untersuchungsgebiet ist zunächst unklar, inwiefern sich lokale oder regionale Zuwachssignale, die an rezenten Beständen festgestellt wurden, erhalten, wenn die rezenten Referenzchronologien mittels historischer Serien erweitert werden. Aus diesem Grund wurde ein weiteres Verfahren (Artikel III) entwickelt. Der vorgestellte Simulationsansatz erlaubt es nun, das zu erwartende Kontaminationsrisiko für Referenzchronologien zu quantifizieren und Schwellenwerte aufgrund der rezenten Datengrundlage transparent herzuleiten. Mit diesem Ansatz kann intersubjektiv nachvollziehbar untersucht werden, ob sich in einem Gebiet anhand rezenter Bestandeschronologien Signalunterschiede abzeichnen und ob diese für die Erstellung lokaler, regionaler oder lediglich höhenlagspezifischer Referenzchronologien ausreichen. Anhand der derzeitig verfügbaren Daten (Fallbeispiel, Kap. 6) scheint für das Einzugsgebiet der Holzversorgung der Stadt Zürich lediglich die Erstellung von höhenlagspezifischen Chronologien sinnvoll. Während es gelang, eine zumindest für einen Teil der Tief- und Mittellagenbestände repräsentative Referenz zu erstellen, war die schliesslich erstellte Höhenlagenchronologie relativ dünn belegt (vgl. Kap. 6). Deshalb stösst der Aufbau von Referenzchronologien derzeit bereits bei solch grober räumlicher Auflösung an seine Grenzen.

7.2 Schlussfolgerungen für die weitere Methodenentwicklung

Viele der bisherigen Dendro-Provenancing Studien (Kap. 2.2–2.3) agieren weitgehend blind. Es wird selten abgeklärt, ob in einem Untersuchungsgebiet überhaupt Zuwachssignale zu erwarten sind, die sich voneinander abgrenzen lassen. Genauso bleibt das Kontaminationsrisiko beim Aufbau von Referenzchronologien derzeit meist unbeachtet. Bei einem solchen Vorgehen werden aber implizit höchste Ansprüche an den Datensatz gestellt, die in den seltensten Fällen erfüllt sein dürften (vgl. Discussion/Conclusion, Artikel III). Deshalb ist beim Grossteil der bisher publizierten Arbeiten mit unzuverlässigen bis falschen Herkunftsbestimmungen zu rechnen.

Trotzdem zeigte sich in den methodischen Untersuchungen im Rahmen des Dissertations-Projektes, dass sich die theoretischen Grundannahmen des Dendro-Provenancings und das Konzept der Lokalchronologien anhand rezenter Daten und mittels Simulationsansatz erhärten lassen. Das Dendro-Provenancing wäre somit eine theoretisch gut fundierte Methode. Jedoch sind grosse Forschungsanstrengungen notwendig, um den implizit hohen Ansprüchen der derzeitigen Anwendung gerecht zu werden. Mögliche Weiterentwicklungen der hier präsentierten Evaluationsmethoden wurden bereits in den jeweiligen Kapiteln aufgezeigt. Generell scheint es, dass mit technischen Fortschritten in der Messtechnik bald zusätzliche Variablen (Früh-/Spätholzbreite, Gefässflächensumme usw.) standardmäßig erfasst werden könnten (Arx/Carrer, 2014; Arx et al., 2016; Loader et al., 2019; Wilson et al., 2017). Daraus ergäben sich multivariate Zeitreihen für Jahrringserien, die auf bestandesspezifische Signale zu testen wären. Solche multivariaten Serien würden jedoch auch eine Weiterentwicklung der hier präsentierten Evaluationsmethoden erfordern, die sich relativ komplex gestalten könnte.

7.3 Schlussfolgerungen für die historische und archäologische Forschung

Das historische Fallbeispiel (Kap. 6) illustriert, wie hoch die Ansprüche an einen dendrochronologischen Datensatz sind, anhand dessen Dendro-Provenancing durchgeführt werden soll. Selbst mit den vereinten Datensätzen der Labore Zürich und Bern, die während Jahrzehnten intensiver dendrochronologischer Tätigkeit aufgebaut worden sind, gelang es nur knapp, die als relativ unproblematisch geltenden Höhenlagenchronologien anhand von Einzelserien aufzubauen (Kap. 6.4–6.5).

Von einer regionalen oder gar lokalen Auflösung des Dendro-Provenancings ist die Forschung in der Schweiz, wie wohl auch in vielen anderen Regionen Europas, noch weit entfernt. Es fehlen die umfangreichen Datenbestände, die sich für eine solch hohe räumliche Auflösung als notwendig erwiesen (Artikel III). Besonders mangelt es im Untersuchungsgebiet an geeigneten Gebäudechronologien. Bislang werden bei Untersuchungen meist nicht mehr als fünf Proben pro Gebäude/Bauphase entnommen.¹³ Darunter finden sich oft mehrere Holzarten sowie zahlreiche undatierbare oder kurze Serien. Die Erhöhung der Probenanzahl ist eine Kostenfrage. Es wird

13 Freundliche mündliche Mitteilung Felix Walder, UWAD ZH. Siehe dazu auch: Eissing, 2010.

nicht leicht sein, auftraggebenden Institutionen und Personen, die primär an der Datierung von Objekten interessiert sind, nahezulegen mehr Proben zu finanzieren, als für die Datierung notwendig sind. Ein weiteres Problem ist, dass der Gefügezusammenhang und die lokale Herkunft des verbauten Holzes bislang nicht systematisch dokumentiert bzw. anhand historischer Quellen überprüft wurde. Die Sensibilisierung für die Frage der Holzherkunft hat auch unter Dendrochronologen noch nicht ausreichend stattgefunden.

Als generelle Empfehlungen für künftige Dendro-Provenancing Untersuchungen lassen sich folgende Punkte festhalten:

- 1) Der Aufbau von Gebäudechronologien sollte, wo immer möglich, durch historische Recherchen zur Holzherkunft und durch eine gefügekundliche Dokumentation der Probenentnahmen begleitet sein.
- 2) Optimalerweise wäre die Entnahme von Proben solange fortzusetzen, bis mindestens sechs datierte Proben mit >50 Jahrringen derselben Holzart pro Gebäude bzw. Bauphase vorhanden sind.
- 3) Eine Evaluation der Datengrundlage mittels der in Artikel I und III präsentierten oder noch weiter entwickelter Methoden ist zwingend dem Dendro-Provenancing voranzustellen.
- 4) Es sollte versucht werden langfristige und umfangreiche, interdisziplinäre Projekte auf nationaler Ebene aufzugleisen. Ohne grossen personellen und finanziellen Aufwand ist keine baldige Verbesserung der Datengrundlage zu erwarten. Auch die Entwicklung von Methoden zur Einbindung multivariater Jahrringserien wird ohne Grossprojekte nicht in absehbarer Zeit möglich sein.

Für die Prähistorische Archäologie scheint das Dendro-Provenancing bis auf Weiteres in der Schweiz kaum nutzbar zu sein. Die Resultate bereits durchgeföhrter Untersuchungen (z. B. in den Niederlanden Jansma et al. (2014); Visser (2015); Domínguez-Delmás et al. (2014)) an archäologischem Material sollten vor dem Hintergrund der hier gemachten Erkenntnisse reflektiert werden. In der Schweiz liessen sich wahrscheinlich allein für die Eiche (Stiel- und/oder Steineiche) Referenzchronologien bis in die Jungsteinzeit aufbauen. Die Eiche kommt auf Trocken- (v. a. Steineiche) wie auf Feuchtstandorten (v. a. Stieleiche) vor. Deshalb ist zu erwarten, dass sich die Wachstumssignale nach Standort unterscheiden. Eine Klassifikation nach Standort wäre folglich ein Forschungsdesiderat. Es ist jedoch unwahrscheinlich, dass sich ausgehend von rezenten Feucht- und Trockenstandortschronologien durchgehende Referenzen bis in die prähistorische Zeit für beide Standorte erstellen lassen. Da das Holz aus den Feuchtbodensiedlungen der Bronze- und Jungsteinzeit

wohl primär in Auenstandorten geschlagen wurde, sind Trockenstandorte unterrepräsentiert (Billamboz, 2014; Bleicher, 2014; Gut, 2013).

7.4 Schlusswort

Das präsentierte Projekt verdeutlicht, wie von einer kulturhistorisch-geisteswissenschaftlichen bzw. archäologischen Fragestellung ausgehend die naturwissenschaftliche Methodenentwicklung inspiriert werden kann. Die Archäologie ist aufgrund ihrer Fragestellungen und ihres Forschungsgegenstandes eine Geisteswissenschaft. Es ist jedoch kaum möglich, dem seit längerem formulierten Ruf nach mehr "Interdisziplinarität" in der Archäologie nachzukommen (Bleicher, 2009, 13 – 16; Haidle, 1998), ohne die Grundlagen für eine Kommunikation zwischen Natur- und Geisteswissenschaften zu schaffen. Nebst Erkenntnistheorie und einer breiten, archäologischen Methodenkompetenz gehört dazu – angesichts des wachsenden Umfangs digital verfügbarer bzw. digitalisierbarer Daten – eine mathematisch-statistische Grundausbildung sowie die Kenntnis einer etablierten und der jeweiligen Anwendung angemessenen Programmiersprache. Am Archäologischen Institut der Universität Zürich scheinen mit der Etablierung eines entsprechenden Minor-Studienprogrammes "Interdisziplinäre Archäologische Wissenschaften" bzw. des Moduls "Computer Applications" konkrete Anstrengungen zur Stärkung dieser Ausbildungsschwerpunkte unternommen worden zu sein. Es ist sehr zu hoffen, dass aus diesen Impulsen weitere interdisziplinäre archäologische Arbeiten hervorgehen werden.

8 Anhang

8.1 Tabellennachtrag zu Artikel I

Tab. A1: Vergleich zwischen der k-NN Performanz bei Verwendung von Korrelationskoeffizienten (alle Ähnlichkeitsmasse mit "r") bzw. t-Werten (alle Ähnlichkeitsmasse mit "t"). Die Masse mit t-Werten als Vergleichsrechnung erreichten durchwegs leicht höhere Werte bezüglich der verschiedenen Performanz Indikatoren. Zur Auflösung der Abkürzungen und Methode siehe: Artikel I, Table 3 bzw. Section 2.

	rAR	tAR	rBP	tBP	rHO	tHO	rDIFF	tDIFF	rSPL10yrs	tSPL10yrs
opening _{ratio}	-none-	0.8	-none-	0.2	-none-	0.8	-none-	0.8	-none-	0.8
on-site–NN _{ratio}	17.8	18.5	16.9	17.9	18.2	18.7	17.9	18.6	17.7	18.4
potential _{ratio}	86.6	87.5	85.5	86.7	87.2	88.2	86.7	87.6	86.8	87.6
CSR	40.4	40.6	34.4	36.0	36.7	38.2	38.6	39.5	36.9	36.7
rating score	36.2	36.9	34.2	35.2	35.5	36.5	35.8	36.6	35.3	35.9
CSR _{open}	-none-	78.5	-none-	76.3	-none-	79.2	-none-	75.9	-none-	75.7

8.2 Exkurs zur statistischen Modellierung der Höhenlagenherkunft anhand von Zuwachsvariablen

Der Einfluss der Höhenlage auf das Wachstum der Fichte ist seit längerem bekannt (siehe Artikel I, Introduction). Dittmar et al. (2012) stellten ein nichtlineares Regressionsmodell vor, mit welchem die Vorhersage der Höhe über dem Meeresspiegel für historische Zuwachsserien in Bayern ermöglicht wurde. Das Untersuchungsgebiet von Dittmar et al. (2012) ist dem hier Untersuchten sehr ähnlich. Gleichzeitig wäre es wünschenswert, die in Kapitel 6 gemachten Höhenlagenzuweisungen mittels einer Modellierung analog derjenigen von Dittmar et al. (2012) zu überprüfen. Im folgenden Exkurs soll knapp auf die im Rahmen der hier vorliegenden Arbeit verfolgten Ansätze eingegangen werden. Diese sind keinesfalls erschöpfend und die gemachten Analysen haben eher explorativen Charakter.

In der Hoffnung innert kurzer Zeit belastbare Resultate zu erhalten, wurden die Höhen über dem Meeresspiegel für alle verfügbaren, rezenten Fichten¹⁴ mittels der von Dittmar et al. (2012) publizierten Formel vorhergesagt. In einem zweiten Schritt wurden eigene multivariate Regressionsmodelle sowie ein *General Additive Model* (GAM) erstellt und verglichen. Diese Modellierungsversuche begannen mit einem

¹⁴ Um die Datengrundlage für die Modellierung zu optimieren, wurden alle Zuwachsseries von Artikel I, Artikel II sowie die Berner Daten, insgesamt 522 Fichten, in die folgenden Analysen integriert. Die kurzen Serien aus Artikel II konnten hier ebenfalls integriert werden, weil für die Modellierungen Jahrringparameter verwendet wurden, die keine Mindestanzahl Ringe erfordern.

nichtlinearen Regressionsmodell, das exakt demjenigen von Dittmar et al. (2012) entsprach, aber an die hier untersuchten Daten angepasst wurde.¹⁵ Die anschliessend angestellten, zusätzlichen Modellierungen basierten auf den erklärenden Variablen *mean sensitivity* (ein dendrochronologisches Standardmass für die Bewegtheit der Zuwachsserie, siehe: Fritts, 1976, 257–258) des ersten autoregressiven Koeffizienten, der maximalen und minimalen Jahrringbreite sowie der Standardabweichung und dem Mittelwert der Zuwachsserien. Bei der Modellierung bzw. Modellselektion wurde gemäss Empfehlungen und Methoden aus Standardlehrbüchern vorgegangen (Crawley, 2015; James et al., 2017; Kuhn/Johnson, 2013). Beim linearen Regressionsmodell sowie dem GAM begann die Modellierung jeweils mit einem sogenannten gesättigten Modell, welches alle in Frage kommenden Zuwachsvariablen integrierte. Das gesättigte Modell wurde anschliessend schrittweise vereinfacht. Zur Modellselektion wurde Akaikes Informations-Kriterium beigezogen (Akaike, 1998). Da sich ausgeprägte nicht-lineare Beziehungen zwischen den erklärenden Variablen abzeichneten, wurden zudem Entscheidungsbaummodelle erstellt, die lediglich die beiden Klassen “Hochlage” bzw. “Tief-/Mittlage” anhand der Variablen Standardabweichung und Mittelwert der Zuwachsserien vorhersagten (Therneau/Atkinson, 2019). Mittels 5-facher Kreuzvalidation (*5-fold cross-validation*) wurde schliesslich die Vorhersageperformanz¹⁶ aller Modelle evaluiert (Kuhn, 2019).

Die Modellierung analog der Formel aus Dittmar et al. (2012) wies eine derart hohe Ungenauigkeit auf (RSME: 452 m), dass die Vorhersagen der Höhen für das hier untersuchte Gebiet unbrauchbar waren. Der Vergleich der Vorhersageperformanz verschiedener Modellierungsansätze in der 5-fachen Kreuzvalidation zeigte schliesslich, dass Entscheidungsbäume tendenziell besser an die hier untersuchten

¹⁵ Das heisst, für das Modell wurden dieselbe Modellformel und dieselben erklärenden Variablen verwendet. Die Optimierung der Modellparameter fand jedoch aufgrund der hier untersuchten Daten statt. Dadurch sollte das Modell besser an die hiesigen Daten angepasst werden und somit genauere Vorhersagen für die Zuwachsserien im Untersuchungsgebiet liefern.

¹⁶ Als Performanzmasse wurden für die Modelle mit einer kontinuierlichen Zielvariable (die Höhe in Metern über Meer) der mittlere quadratische Fehler (*Root-Mean-Square Error*, RSME) sowie das Bestimmtheitsmass (R^2) der Kreuzvalidierung verwendet (Kuhn and Johnson, 2013). Für die Modellierung kategorialer Zielvariablen (“Höhenlage” bzw. “Tief-/Mittellage”) wurde die in der Kreuzvalidierung erzielte Treffergenauigkeit (*Accuracy*) sowie Cohen’s Kappa berechnet (Kuhn and Johnson, 2013). Cohen’s Kappa ist ein Mass, welches die beobachtete Treffergenauigkeit mit der erwarteten Treffergenauigkeit vergleicht (Cohen, 1968, 1960). Kappa ist am höchsten, wenn die erwartete Treffergenauigkeit Null (bzw. aufgrund von Zufall kein Datenpunkt korrekt klassifiziert werden konnte) und die beobachtete Treffergenauigkeit 1 ist (also mit dem erstellten Modell alle Daten korrekt klassifiziert wurden).

Daten angepasst werden konnten (Tab. 4). Das favorisierte Modell war folglich ein Entscheidungsbaummodell, welches lediglich die Variablen Jahrringbreiten-Standardabweichung und Jahrringbreiten-Mittelwert zur Modellierung berücksichtigte (vgl. Tab. 5, Abb. 6).

*Tab. 4: Performanz der Modelle, welche die Höhe über Meer in Metern für Zuwachs-
serien als kontinuierliche Zahlenwerte vorhersagen. Kürzel: "nls" = Nicht-lineares
Modell analog zu Dittmar et al. 2012; "lm" = Lineares Modell mit Mean Sensitivity und
erstem autoregressivem Koeffizienten als erklärende Variablen; "gam" = General
Additive Model basierend auf allen im Fliesstext genannten Zuwachsvariablen ausser
der maximalen Jahrringbreite. Die Residuen-Diagnose-Plots der Modelle "nls" und
"lm" deuteten darauf hin, dass die Modellannahmen "gleichbleibende Varianz" und
"Normalverteilung" der Residuen verletzt waren. Die Residuen-Diagnose-Plots für
das "gam"-Modell hingegen waren zufriedenstellend. Alle Modelle wiesen in der 5-
fachen Kreuzvalidation einen hohen mittleren quadratischen Fehler (RMSE) sowie
einen niedrigen R^2 auf. Die hohen RMSE der Kreuzvalidation weisen darauf hin, dass
die Vorhersagen für neue Daten (Testdaten), die
nicht bei der Modellierung verwendet wurden, un-
genau sind (bzw. eine hohe Streuung aufweisen).
Die niedrigen R^2 deuten zudem darauf hin, dass die
Modelle die Varianz in den Testdaten nur zu einem
kleinen Teil erklären (approximativ zwischen 13%
und 25% erklärte Varianz).*

	RMSE	R^2
nls	366.66 m	0.13
lm	359.26 m	0.16
gam	341.39 m	0.25

Tab. 5: Die Entscheidungsbaummodelle wurden mittels Recusive Partitioning in R mit Package "rpart" erstellt (Therneau, 2019). Die Treffergenauigkeit (Accuracy) ist in allen Modellen über 70%. Die mässig hohen Werte für Cohen's Kappa deuten jedoch darauf hin, dass die erwartete Treffergenauigkeit (durch Zufall) zumindest für eine der Klassen ("Hochlagen" oder "Tief-/Mittellagen") relativ hoch ist (siehe auch Fussnote 16 sowie Abb. 7). In Abb. 7 ist klar ersichtlich, dass die Vorhersage der Klasse "Hochlagen" durch das Modell "rpart" weniger zuverlässig funktioniert als die der Klasse "Tief-/Mittellagen". Es wurde vermutet, dass dies eine Konsequenz der geringeren relativen Häufigkeit von rezenten Serien aus Hochlagen im untersuchten Datensatz sein könnte. Deshalb wurde mittels der Verfahren Over- bzw. Undersampling versucht diese Unterschiede in der relativen Häufigkeit der beiden Klassen auszugleichen. Entsprechend wurden zwei weitere Modelle erstellt ("rpart_o" verwendet Oversampling, "rpart_u" Undersampling). Die Performanzmasse aller drei rpart-Modell sind vergleichbar. Das Problem der hohen Unzuverlässigkeit bei der Vorhersage der Hochlagenserien konnte somit nicht mittels Over- bzw. Undersampling gelöst werden. Während beim Modell "rpart" vermehrt Fehlklassifikationen bei den Hochlagenserien auftreten, geschehen Fehlklassifikation bei den Modellen "rpart_o" und "rpart_u" vornehmlich bei den Tief- und Mittellagenserien (siehe Abb. 7).

	Accuracy	Kappa
rpart	0.79	0.44
rpart_o	0.72	0.38
rpart_u	0.76	0.42

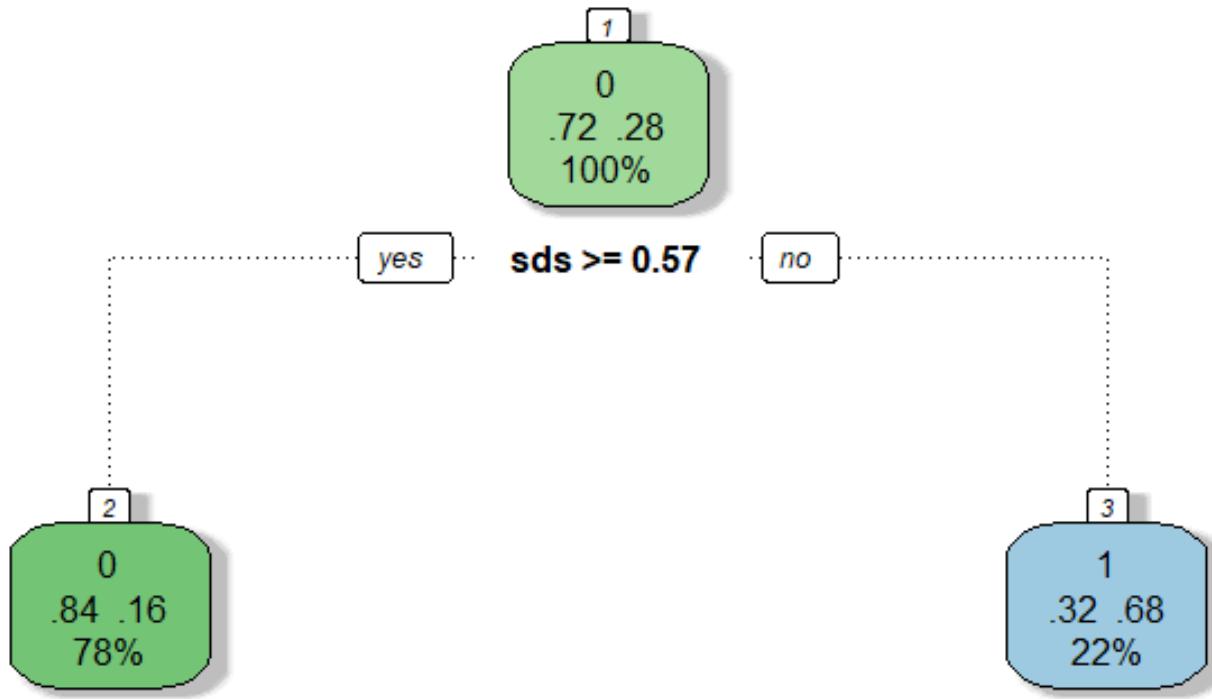


Abb. 6: *Finales Entscheidungsbaummodell*. Der Baum weist eine sehr reduzierte Struktur auf und basiert lediglich auf der Jahrringbreitenstandardabweichung (sds). Serien, für welche die Standardabweichung ≥ 0.57 mm ist, werden als "Tief- und Mittellagen" (Klasse "0") und der Rest als "Hochlagen" (Klasse "1") klassifiziert. Prozentwerte geben an, welcher Anteil an Serien des Gesamtdatensatzes im betreffenden Ast gelandet ist. Die Anteile über den Prozentangaben zeigen die Verteilung der realen Klassen im betreffenden Ast an. Z. B. im Schluss-Ast der Klasse "0", in grün, bedeutet der Wert "0.84", dass 84% der als Klasse "0" klassifizierten Serien tatsächlich aus Tief- und Mittellagen stammen. Der Wert "0.16" zeigt an, dass 16% der Serien mit Klasse "0" eigentlich aus Hochlagenbeständen kommen.

Der Vergleich der Höhenlagenzuweisungen dieses Modells mit den Zuweisungen, die anhand des Signalverlaufs in Kapitel 6 gemacht wurden, zeigte nur geringe Übereinstimmungen. Zwar wurden 25 von 36 Serien, die mittels des Signalverlauf-Ansatzes als Hochlagenserien klassifiziert wurden, ebenfalls vom Entscheidungsbaummodell als Hochlagenserien klassifiziert, doch wurden zudem 90 Serien den Hochlagen zugewiesen, die im Signalverlauf-Ansatz nicht klassifiziert worden waren. Tendenziell wurden mit dem Entscheidungsbaummodell also merklich mehr Serien als Hochlagenserien klassifiziert. Jedoch wies das Modell eine Schwäche bei der Vorhersage der Klasse "Hochlage" auf: Eine relativ hohe Anzahl "Hochlagen"-Serien wurden infolgedessen falsch als "Tief- und Mittellagen"-Serien klassifiziert (Abb. 7). Ein Grund für diese Fehlklassifikationen liegt wahrscheinlich darin, dass die verwendeten Variablen (Mittelwert und Standardabweichung) eher mit der Geschwindigkeit des Baumwachstums als mit der Höhe zusammenhängen. Zwar wachsen

Bäume in Hochlagen aufgrund des raueren Klimas tendenziell langsamer, doch können auch in Tief- und Mittellagen Bäume, die im Schatten anderer oder unter Nährstoff- sowie Wasserknappheit wachsen, ebenfalls langsam wachsen. Umgekehrt kann auch in Hochlagen an einem eher geschützten Standort die Wachstumsgeschwindigkeit hoch sein. Das heisst, ein generell langsam wachsender Baum wird als Höhenlagenbaum klassifiziert, obwohl es sich beim Wuchssignal dieses Baumes nicht um ein Höhenlagensignal handelt, und ein tendenziell schnell wachsender Baum wird als Tief- oder Mittellagenbaum klassifiziert, obwohl es sich beim Wuchssignal dieses Baumes nicht um ein Tief- oder Mittellagensignal handelt. Von derselben Problematik ist auch die von Dittmar et al. (2012) verwendete Variable, die Mean Sensitivity, betroffen, die überdies von Bunn et al. (2013) als zweifelhafte Variable für die Beschreibung der statistischen Eigenschaften von Jahrringserien eingestuft wurde.

		rpart		rpart_o		rpart_u	
		Vorhersage		Vorhersage		Vorhersage	
		MittelTief	Hoch	MittelTief	Hoch	MittelTief	Hoch
Reale Klasse	MittelTief	340	37	305	72	337	40
	Hoch	67	78	42	103	65	80

Abb. 7: Konfusionsmatrizen der verschiedenen “rpart”-Modelle (Entscheidungsbaummodelle). Kürzel: “rpart” = Entscheidungsbaummodell basierend auf den Variablen Jahrringbreitenstandardabweichung und -mittelwert (der Rohwertserien); “rpart_o” = Entscheidungsbaummodell mit denselben Variablen wie “rpart”, aber mit Oversampling (Kuhn, 2019); “rpart_u” = Entscheidungsbaummodell mit denselben Variablen wie “rpart”, aber mit Undersampling (Kuhn, 2019). Die Spalteneinträge entsprechen der Anzahl Serien, welche je Klasse vorhergesagt wurden. Das Aufsummieren der Zeileneinträge ergibt die tatsächliche Anzahl Serien pro Klasse im Datensatz. Es ist erkennbar, dass beim Modell “rpart” nur 37 Tief- und Mittellagenserien falsch der Klasse “Hoch” zugewiesen wurden, während dieselbe Art der Falschzuweisung bei den beiden anderen Modellen höher liegt (72 bzw. 40). Jedoch wurden von den Modellen “rpart_o” und “rpart_u” weniger Hochlagenserien falsch als “Mittel-Tief” klassifiziert. Vereinfacht ausgedrückt erzeugen die Modelle mit Oversampling bzw. Undersampling zwar weniger Falschklassifikationen unter den Hochlagenserien, doch gleichzeitig werden mehr eigentlich aus den Tief-/Mittellagen stammende, rezente Serien als Hochlagenserien klassifiziert. Unter dem Strich wurde also mit den Techniken Oversampling bzw. Undersampling hier keine Modellverbesserung erzielt, was sich auch in der etwa gleichbleibenden Performanz in der Kreuzvalidation zeigte (vgl. Tab. 5).

Die entwickelten Regressions- und GAM-Modelle sowie das bereits von Dittmar et al. (2012) publizierte Modell waren entweder sehr schlecht an die hier untersuchten Daten anzupassen und/oder wiesen eine hohe Ungenauigkeit auf. Die hier präsentierten Modelle repräsentieren nur eine Auswahl denkbarer Modellierungsansätze. Es mögen durchaus besser geeignete Ansätze existieren. Doch es muss bedacht werden, dass die bislang verwendeten, erklärenden Variablen nicht allein von der Höhenlage abhängig sind. Das Wuchssignal hingegen wurde im Untersuchungsgebiet als höhenlagenspezifisch identifiziert (Artikel I und III). Deshalb kann solchen Klassifikationen, die auf dem Wuchssignal basieren, hier eine grössere Plausibilität attestiert werden als Modellierungen wie den erwähnten Regressions-, GAM- bzw. Entscheidungsbaummodellen, für deren Erstellung keine das Wuchssignal charakterisierende Variablen berücksichtigt worden sind.

9 Literaturverzeichnis

Abanda et al., 2019

A. Abanda/U. Mori/J. A. Lozano, A review on distance based time series classification. *Data Mining and Knowledge Discovery* 33, 2019, 378–412.

Akaike, 1998

H. Akaike, Information theory and an extension of the maximum likelihood principle. In: E. Parzen/K. Tanabe/G. Kitagawa (Hrsg.), *Selected papers of Hirotugu Akaike*. Springer Series in Statistics (New York 1998) 199–213.

Akhmetzyanov et al., 2019

L. Akhmetzyanov/A. Buras/U. Sass-Klaassen et al., Multi-variable approach pinpoints origin of oak wood with higher precision. *Journal of Biogeography* 46, 6, 2019, 1163–1177.

Arx/Carrer, 2014

G. von Arx/M. Carrer, ROXAS – A new tool to build centuries-long tracheid-lumen chronologies in conifers. *Dendrochronologia* 32, 3, 2014, 290–293.

Arx et al., 2016

G. von Arx/A. Crivellaro/A. L. Prendin et al., Quantitative wood anatomy. Practical guidelines. *Frontiers in Plant Science* 7, 2016, 781.

Babst et al., 2018

F. Babst/P. Bodesheim/N. Charney et al., When tree rings go global. Challenges and opportunities for retro- and prospective insight. *Quaternary Science Reviews* 197, 2018, 1–20.

Baillie/Pilcher, 1973

M. G. L. Baillie/J. R. Pilcher, A simple cross-dating program for tree-ring research. *Tree-Ring Bulletin* 33, 1973, 7–14.

Bauch/Eckstein, 1970

J. Bauch/D. Eckstein, Dendrochronological dating of oak panels of Dutch seventeenth-century paintings. *Studies in Conservation* 15, 1, 1970, 45–50.

Bernabei/Bontadi, 2011

M. Bernabei/J. Bontadi, Determining the resonance wood provenance of stringed instruments from the Cherubini Conservatory Collection in Florence, Italy. *Journal of Cultural Heritage* 12, 2, 2011, 196–204.

Billamboz, 2014

A. Billamboz, Regional patterns of settlement and woodland developments. *Dendroarchaeology in the Neolithic pile-dwellings on Lake Constance (Germany)*. *The Holocene* 24, 10, 2014, 1278–1287.

Bitterli, 2009

D. Bitterli, Holzhandel zwischen Zürich und dem Kloster Einsiedeln. In: Äkte, Amtsleute, Archivare. Zürich und das Kloster Einsiedeln. Mitteilungen der Antiquarischen Gesellschaft in Zürich 76 (2009) 59–69.

Bleicher, 2009

N. Bleicher, Altes Holz in neuem Licht. Archäologische und dendrochronologische Untersuchungen an spätneolithischen Feuchtbodensiedlungen in Oberschwaben. Materialhefte zur Archäologie in Baden-Württemberg 83 (Stuttgart 2009).

Bleicher, 2014

N. Bleicher, Pastoralism, Landscape and woodland use in the Late Neolithic in Southern Germany. In: A. G. Heiss/F. Retamero/I. Schjellerup et al. (Hrsg.), Agricultural and Pastoral Landscapes in Pre-Industrial Society (2014) 239–250.

Boswijk/Fowler, 2019

G. Boswijk/A. M. Fowler, Dendroprovenancing. A preliminary assessment of potential to geo-locate kauri timbers in Northern New Zealand. *Dendrochronologia* 57, 2019, 125611, 1–13.

Bridge, 2000

M. Bridge, Can dendrochronology be used to indicate the source of oak within Britain? *Vernacular Architecture* 31, 1, 2000, 67–72.

Bridge, 2012

M. Bridge, Locating the origins of wood resources. A review of dendroprovenancing. *Journal of Archaeological Science* 39, 8, 2012, 2828–2834.

Bridge/Fowler, 2019

M. C. Bridge/A. M. Fowler, A new way of looking at dendroprovenancing. Spatial field correlations of residuals. *Dendrochronologia* 57, 2019, 125627, 1–10.

Bunn et al., 2013

A. G. Bunn/E. Jansma/M. Korpela et al., Using simulations and data to evaluate mean sensitivity (ζ) as a useful statistic in dendrochronology. *Dendrochronologia* 31, 3, 2013, 250–254.

Bürgi/Schuler, 2001

M. Bürgi/A. Schuler, Topography, technology and demand for timber. Entrepreneurs and foresters shaping the landscape in a Swiss alpine valley. *Forest History Today*, 2002, Fall 2002, 20–24.

Cohen, 1960

J. Cohen, A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 20, 1, 1960, 37–46.

Cohen, 1968

J. Cohen, Weighted kappa. Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin* 70, 4, 1968, 213–220.

Conradi et al., unpubliziert

C. Conradi/K. Knüsel/Y. Shi, Classification of tree ring-width time series into different regions.
Supervised by Nina Roth. Statistics Lab Report, ETH Zurich, Department of Mathematics,
Seminar for Statistics, 23.06.2016.

Cook, 1985

E. R. Cook, A time series approach to tree ring standardization (University of Arizona: Dissertation, 1985).

Cook/Kairiukstis, 1990

E. R. Cook/L. A. Kairiukstis (Hrsg.), Methods of dendrochronology (Dordrecht 1990).

Crawley, 2015

M. J. Crawley, Statistics. An introduction using R (Chichester 2015).

Daly, 2007a

A. Daly, Timber, trade and tree-rings. A dendrochronological analysis of structural oak timber in Northern Europe, c. AD 1000 to c. AD 1650 (University of Southern Denmark: Dissertation, 1985).

Daly, 2007b

A. Daly, The Karschau Ship, Schleswig-Holstein. Dendrochronological results and timber provenance. *International Journal of Nautical Archaeology* 36, 1, 2007, 155–166.

Daly, 2011

A. Daly, Dendro-geography. Mapping the Northern European historic timber trade. In: P. Fraiture (Hrsg.), Tree rings, art and archaeology (2011) 107–123.

Daly/Nymoen, 2008

A. Daly/P. Nymoen, The Bøle Ship, Skien, Norway. Research history, dendrochronology and provenance. *International Journal of Nautical Archaeology* 37, 1, 2008, 153–170.

Deguilloux et al., 2006

M. F. Deguilloux/L. Bertel/A. Celant et al., Genetic analysis of archaeological wood remains. First results and prospects. *Journal of Archaeological Science* 33, 9, 2006, 1216–1227.

Diaz et al., 2011

H. F. Diaz/M. K. Hughes/T. W. Swetnam, Dendroclimatology. Progress and prospects. *Developments in Paleoenvironmental Research* 11 (Dordrecht 2011).

Dittmar et al., 2012

C. Dittmar/T. Eissing/A. Rothe, Elevation-specific tree-ring chronologies of Norway spruce and silver fir in Southern Germany. *Dendrochronologia* 30, 2, 2012, 73–83.

Domínguez-Delmás et al., 2013

M. Domínguez-Delmás/N. Nayling/T. Ważny et al., Dendrochronological dating and provenancing of timbers from the Arade 1 shipwreck, Portugal. *International Journal of Nautical Archaeology* 42, 1, 2013, 118–136.

Domínguez-Delmás et al., 2014

M. Domínguez-Delmás/M. Driessen/I. García-González et al., Long-distance oak supply in mid-2nd century AD revealed. The case of a Roman harbour (Voorburg-Arentsburg) in the Netherlands. *Journal of Archaeological Science* 41, 2014, 642–654.

Domínguez-Delmás et al., 2015

M. Domínguez-Delmás/R. Alejano-Monge/S. van Daalen et al., Tree-rings, forest history and cultural heritage. Current state and future prospects of dendroarchaeology in the Iberian Peninsula. *Journal of Archaeological Science* 57, 2015, 180–196.

Domínguez-Delmás et al., 2018

M. Domínguez-Delmás/S. van Daalen/R. Alejano-Monge et al., Timber resources, transport and woodworking techniques in post-medieval Andalusia (Spain). Insights from dendroarchaeological research on historic roof structures. *Journal of Archaeological Science* 95, 2018, 64–75.

Drake, 2018

B. L. Drake, Source & Sourceability. Towards a probabilistic framework for dendroprovenance based on hypothesis testing and Bayesian inference. *Dendrochronologia* 47, 2018, 38–47.

Eckstein/Wrobel, 2007

D. Eckstein/S. Wrobel, Dendrochronological proof of origin of historic timber. Retrospect and perspectives. In: K. Haneca/Verheyden A./H. Beeckman et al. (Hrsg.), TRACE - Tree Rings in Archaeology, Climatology and Ecology 5. Proceedings of the Dendrosymposium 2006, April 20th - 22nd 2006, Tervuren, Belgium. Schriften der Forschungszentrums Jülich, Reihe Umwelt (2007) 8–20.

Eissing, 2007

T. Eissing, Zum Problem der Interpretation dendrochronologischer Datierung bei Flossholz am Beispiel Bamberg. In: Arbeitskreis für Hausforschung (Hrsg.), Spuren der Nutzung in historischen Bauten. *Jahrbuch für Hausforschung* 54 (2007) 23–36.

Eissing/Dittmar, 2011

T. Eissing/C. Dittmar, Timber transport and dendro-provenancing in Thuringia and Bavaria. In: P. Fraiture (Hrsg.), *Tree rings, art and archaeology* (2011) 137–149.

Eissing, 2010

T. Eissing, Holzversorgung, Holzverbrauch, Holzherkunft am Beispiel Mitteldeutschlands. In: A. Dix/W. Schenk (Hrsg.), Konsum und Kulturlandschaft (2010) 87–106.

Ellenberg, 1996

H. Ellenberg, Vegetation Mitteleuropas mit den Alpen in ökologischer, dynamischer und historischer Sicht (Stuttgart 1996).

English et al., 2001

N. B. English/J. L. Betancourt/J. S. Dean et al., Strontium isotopes reveal distant sources of architectural timber in Chaco Canyon, New Mexico. *Proceedings of the National Academy of Sciences of the United States of America* 98, 21, 2001, 11891–11896.

Fang et al., 2015

K. Fang/D. Chen/X. Gou et al., Influence of non-climatic factors on the relationships between tree growth and climate over the Chinese Loess Plateau. *Global and Planetary Change* 132, 2015, 54–63.

Fowler/Bridge, 2015

A. Fowler/M. Bridge, Mining the British Isles oak tree-ring data set. Part A: Rationale, data, software, and proof of concept. *Dendrochronologia* 35, 2015, 24–33.

Fowler/Bridge, 2017

A. M. Fowler/M. C. Bridge, Empirically-determined statistical significance of the Baillie and Pilcher (1973) t statistic for British Isles oak. *Dendrochronologia* 42, 2017, 51–55.

Fowler et al., 2017

A. M. Fowler/M. C. Bridge/G. Boswijk, An empirical resampling method for determining optimal high-pass filters used in correlation-based tree-ring crossdating. *Dendrochronologia* 44, 2017, 84–93.

Fraiture, 2009

P. Fraiture, Contribution of dendrochronology to understanding of wood procurement sources for panel paintings in the former Southern Netherlands from 1450 AD to 1650 AD. *Dendrochronologia* 27, 2, 2009, 95–111.

Fritts, 1976

H. C. Fritts, Tree rings and climate (London 1976).

García-González, 2008

I. García-González, Comparison of different distance measures for cluster analysis of tree-ring series. *Tree-Ring Research* 64, 1, 2008, 27–37.

Grabner et al., 2018

M. Grabner/G. Buchinger/M. Jeitler, Stories about building history told by wooden elements. Case studies from Eastern Austria. International Journal of Architectural Heritage 12, 2, 2018, 178–194.

Grossmann, 1972

H. Grossmann, Flösserei und Holzhandel aus den Schweizer Bergen bis zum Ende des 19. Jahrhunderts. Mitteilungen der Antiquariatischen Gesellschaft in Zürich 46 (Zürich 1972).

Gut, 2013

U. Gut, Spätbronzezeitliche Besiedlungsdynamik und Waldwirtschaft am untersten Zürichsee. Eine dendroarchäologische Perspektive zur Frage der Besiedlung der Strandplatten am untersten Seebecken und der Bewirtschaftung des angrenzenden Hinterlandes (Universität Zürich: Lizentiat, 2013).

Haidle, 1998

M. N. Haidle, Interdisziplinarität in der Archäologie: eine Notwendigkeit?! Archäologische Informationen 21, 1, 9–20.

Hajj et al., 2017

F. Hajj/A. Poszwa/J. Bouchez et al., Radiogenic and “stable” strontium isotopes in provenance studies. A review and first results on archaeological wood from shipwrecks. Journal of Archaeological Science 86, 2017, 24–49.

Haneca et al., 2005

K. Haneca/T. Wasny/J. van Acker et al., Provenancing Baltic timber from art historical objects. Success and limitations. Journal of Archaeological Science 32, 2005, 261–271.

Hollstein, 1980

E. Hollstein, Mitteleuropäische Eichenchronologie. Trierer dendrochronologische Forschungen zur Archäologie und Kunstgeschichte 11 (Mainz am Rhein 1980).

Hughes, 1982

M. K. Hughes (Hrsg.), Climate from tree rings (Cambridge 1982).

James et al., 2017

G. James/D. Witten/T. Hastie et al., An introduction to statistical learning - with applications in R. Springer Texts in Statistics (New York 2017).

Jansma et al., 2014

E. Jansma/K. Haneca/M. Kosian, A dendrochronological reassessment of three Roman boats from Utrecht (the Netherlands). Evidence of inland navigation between the lower-Scheldt region in *Gallia Belgica* and the *limes* of *Germania inferior*. Journal of Archaeological Science 50, 2014, 1–13.

Kagawa/Leavitt, 2010

A. Kagawa/S. W. Leavitt, Stable carbon isotopes of tree rings as a tool to pinpoint the geographic origin of timber. *Journal of Wood Science* 56, 3, 2010, 175–183.

Kaya/Gündüz-Öğüdücü, 2015

H. Kaya/Ş. Gündüz-Öğüdücü, A distance based time series classification framework. *Information Systems* 51, 2015, 27–42.

King et al., 2013

G. M. King/F. Gugerli/P. Fonti et al., Tree growth response along an elevational gradient. Climate or genetics? *Oecologia* 173, 4, 2013, 1587–1600.

Kuhn/Johnson, 2013

M. Kuhn/K. Johnson, *Applied predictive modeling* (New York 2013).

Kuniholm et al., 2007

P. I. Kuniholm/C. B. Griggs/M. W. Newton, Evidence for early timber trade in the Mediterranean. In: K. Belke (Hrsg.), *Byzantina Mediterranea. Festschrift für Johannes Koder zum 65. Geburtstag* (2007) 365–384.

Loader et al., 2019

N. J. Loader/D. McCarroll/D. Miles et al., Tree ring dating using oxygen isotopes. A master chronology for Central England. *Journal of Quaternary Science* 100, 6529, 2019, 374.

Maria, 1997

A. Maria, Introduction to modelling and simulation. In: S. Andradóttir/K. J. Healy/D. H. Withers et al. (Hrsg.), *Winter Simulation Conference Proceedings* (1997) 7–13.

Meiggs, 1990

R. Meiggs, *Trees and timber in the ancient Mediterranean world* (London 1998).

Pott, 1993

R. Pott, *Farbatlas Waldlandschaften. Ausgewählte Waldtypen und Waldgesellschaften unter dem Einfluß des Menschen* (Stuttgart 1993).

Rachmayanti et al., 2009

Y. Rachmayanti/L. Leinemann/O. Gailing et al., DNA from processed and unprocessed wood. Factors influencing the isolation success. *Forensic Science International: Genetics* 3, 3, 2009, 185–192.

Rich et al., 2016

S. Rich/S. W. Manning/P. Degryse et al., To put a cedar ship in a bottle. Dendroprovenancing three ancient East Mediterranean watercraft with the $^{87}\text{Sr}/^{86}\text{Sr}$ isotope ratio. *Journal of Archaeological Science: Reports* 9, 2016, 514–521.

Rich, 2013

S. A. Rich, Ship timber as symbol? Dendro-provenancing & contextualizing ancient cedar ship remains from the Eastern Mediterranean / Near East (Katholieke Universiteit Leuven: Dissertation, 2013).

Rodríguez-Trobajo/Domínguez-Delmás, 2015

E. Rodríguez-Trobajo/M. Domínguez-Delmás, Swedish oak, planks and panels. Dendroarchaeological investigations on the 16th century *Evangelistas* altarpiece at Seville Cathedral (Spain). Journal of Archaeological Science 54, 2015, 148–161.

Sass-Klaassen et al., 2008

U. Sass-Klaassen/T. Vernimmen/C. Baittinger, Dendrochronological dating and provenancing of timber used as foundation piles under historic buildings in the Netherlands. International Biodeterioration & Biodegradation 61, 1, 2008, 96–105.

Savva et al., 2006

Y. Savva/J. Oleksyn/P. B. Reich et al., Interannual growth response of Norway spruce to climate along an altitudinal gradient in the Tatra Mountains, Poland. Trees 20, 6, 2006, 735–746.

Schmitt, 2006

I. Schmitt, Ähnlichkeitssuche in Multimedia-Datenbanken. Retrieval, Suchalgorithmen und Anfragebehandlung (München 2006).

Schweingruber, 1983

F. H. Schweingruber, Der Jahrring. Standort, Methodik und Klima in der Dendrochronologie (Bern 1983).

Schweingruber, 1996

F. H. Schweingruber, Tree rings and environment. Dendroecology (Bern 1996).

Speer, 2010

J. H. Speer, Fundamentals of tree-ring research (Tucson 2010).

Traoré et al., 2017

M. Traoré/J. Kaal/A. Martínez Cortizas, Potential of pyrolysis-GC–MS molecular fingerprint as a proxy of modern age Iberian shipwreck wood preservation. Journal of Analytical and Applied Pyrolysis 126, 2017, 1–13.

van Daalen/van der Beek, 2004

S. van Daalen/J. van der Beek, Dendroprovenancing ship's timbers. A pilot study on a Dutch 18th century 'ventjager'. In: E. Jansma/A. Bräuning/H. Gärtner et al. (Hrsg.), TRACE - Tree Rings in Archaeology, Climatology and Ecology 2. Proceedings of the Dendrosymposium 2003, May 1st - 3rd 2003, Utrecht, the Netherlands. Schriften des Forschungszentrums Jülich Reihe Umwelt 44 (2004) 123–130.

Visser, 2015

R. M. Visser, Imperial timber? Dendrochronological evidence for large-scale road building along the Roman *limes* in the Netherlands. *Journal of Archaeological Science* 53, 2015, 243–254.

Ważny, 2002

T. Ważny, Baltic timber in Western Europe. An exciting dendrochronological question. *Dendrochronologia* 20, 3, 2002, 313–320.

Ważny, 2011

T. Ważny, Dendro-provenancing between the Baltic Sea and the East Mediterranean. In: P. Fraiture (Hrsg.), *Tree rings, art and archaeology* (2011) 81–87.

Ważny et al., 2014

T. Ważny/B. Lorentzen/N. Köse et al., Bridging the gaps in tree-ring records. Creating a high-resolution dendrochronological network for Southeastern Europe. *Radiocarbon* 56, 4, 2014, 39–50.

Weisz et al., 1983

L. Weisz/H. Grossmann/E. Krebs et al., *Forstpolitik, Forstverwaltung und Holzversorgung im Alten Zürich. 650 Jahre Zürcherische Forstgeschichte* 1 (Zürich 1983).

Wigley et al., 1987

T.M.L. Wigley/P. D. Jones/K. R. Briffa, Cross-dating methods in dendrochronology. *Journal of Archaeological Science* 14, 1, 1987, 51–64.

Wilson/Hopfmüller, 2001

R. Wilson/M. Hopfmüller, Dendrochronological investigations of Norway spruce along an elevational transect in the Bavarian Forest, Germany. *Dendrochronologia* 19, 1, 2001, 67–79.

Wilson et al., 2017

R. Wilson/D. Wilson/M. Rydval et al., Facilitating tree-ring dating of historic conifer timbers using Blue Intensity. *Journal of Archaeological Science* 78, 2017, 99–111.

Winiwarter/Knoll, 2007

V. Winiwarter/M. Knoll, *Umweltgeschichte* (Köln 2007).

WMO, 1989

World Meteorological Organization (WMO), Calculation of monthly and annual 30 year standard normals. *World Climate Data Programme Series* 10, 1989.

Xing et al., 2010

Z. Xing/J. Pei/E. Keogh, A brief survey on sequence classification. *SIGKDD Explorations Newsletter* 12, 1, 2010, 40–48.

Online Ressourcen

Dubler, 2011

A.-M. Dubler, Flösserei. In: Historisches Lexikon der Schweiz (HLS), (2011), <<https://hls-dhs-dss.ch/de/articles/014055/2009-11-05>> [Stand: 25. September 2019].

Körner/Degen, 2011

M. Körner/B. Degen, Versicherungen. 2.2 Schadenversicherung. In: Historisches Lexikon der Schweiz (HLS), (2011), <<https://hls-dhs-dss.ch/de/articles/014066/2014-01-21>> [Stand: 25. September 2019]

Kuhn, 2019

M. Kuhn, caret: Classification and regression training. R package version 6.0-84. <<https://CRAN.R-project.org/package=caret>> [Stand: 25. September 2019].

Therneau/Atkinson, 2019

T. Therneau/B. Atkinson, rpart: Recursive partitioning and regression trees. R package version 4.1-15. <<https://CRAN.R-project.org/package=rpart>> [Stand: 24. September 2019].

II. Appendix

A. Eigenleistung

Die drei im Rahmen des Projektes verfassten Artikel sind in peer-reviewten Fachzeitschriften entweder bereits publiziert worden oder sind zum Zeitpunkt der Abgabe der Promotionsdokumente zur Publikation in einer solchen Zeitschrift eingereicht worden.

Artikel I und III sind in Allein-Autorenschaft erstellt worden.

Artikel II wurde im Rahmen eines internationalen Netzwerks geschrieben. Ich hatte den Lead bei der Verfassung des Manuskriptes und hatte die Hauptverantwortung für dessen Inhalt. Der Artikel enthält ein *“Author Contribution Statement”*, das hier noch einmal abgedruckt wird:

“The scientific concept and research question of this study was developed in collaboration by A. Buras and U. Gut. The latter also conducted the bulk of the analyses presented here. A. Buras initiated the tree-ring data network. Both together have worked on the submitted manuscript. All other members of the network are listed alphabetically as co-authors, with the exception of Ute Sass-Klaassen, who hosted the first network workshop at Wageningen University and Research. These co-authors contributed to the article by providing tree-ring data sets and valuable feedback at various stages of data analyses as well as commenting on the final manuscript version before submission. We are very grateful for the financial support provided by the Association of Tree-ring Research (ATR), which allowed the implementation of a network-internal statistical workshop.” (Gut et al., 2019, 8)

B. Publikationen

(1) Artikel I

Evaluating the key assumptions underlying dendro-provenancing: How to spruce it up with a scissor plot.

Urs Gut

Dieser Artikel wurde veröffentlicht in:

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



Original article

Evaluating the key assumptions underlying dendro-provenancing: How to spruce it up with a scissor plot

Urs Gut^{a,b,*}^a Universität Zürich, Institut für Archäologie, Fachbereich Prähistorische Archäologie, Karl Schmid-Strasse 4, 8006 Zürich, Switzerland^b ETH Zürich, Department of Environmental System Science, Institute of Terrestrial Ecosystems, Forest Ecology, Universitätstrasse 22, 8092 Zürich, Switzerland

ARTICLE INFO

Keywords:

Dendrochronology
Method
Cross-validation
Nearest neighbor classification
Elevational transect
Proximity measure

ABSTRACT

The pairwise statistical comparison of ring-width series is the basic analysis of dendro-provenancing studies. It is assumed that statistical proximity indicates similar provenance, but this assumption often remains untested. Especially for small areas with high topographic complexity, it is unknown to what extent statistical proximity and geographical provenance are correlated.

In this paper, dendro-provenancing is framed as a search for statistical Nearest Neighbors. The ‘k-Nearest Neighbors leave one-out cross-validation’ process (k-NN) is proposed as a method for validating dendro-provenancing approaches. Furthermore, it allows researchers to consistently compare and evaluate different proximity measures with respect to their suitability for dendro-provenancing. The validation process is demonstrated on a data set of 401 ring-width series of Norway spruce (*Picea abies (L.) H. Karst.*) encompassing 15 sites along elevational gradients in north-eastern Switzerland. Moreover, a new type of plot, the so-called scissor plot, is introduced to visualize the k-NN validation process.

Results indicate that dendro-provenancing depends heavily on differences in between sites high-frequency signal. Mean classification success for the relevant stages of the k-NN (CSR_{open})¹ ranged from 71.8% to 79.2% for the best performing measures. Classification errors occurred mainly between sites at elevations of 1000–1198 m a.s.l. At all other elevations and between different regions of the study area, only moderate differences in classification performance were detected. Thus, the results indicate that dendro-provenancing may be principally feasible even in a small region as studied here.

1. Introduction

Knowing the site of tree growth for timber in archaeological or historical structures and artefacts (e.g., buildings, paintings, ships, etc.) provides crucial information for reconstructing timber trade routes, determining the provenance and authenticity of art historical objects, or estimating forcing factors on tree growth (Wazny, 2002; Eissing and Dittmar, 2011; Jansma et al., 2014; Hellmann et al., 2017). The bulk of this so-called *dendro-provenancing* relies on pair-wise comparisons between ring-width series of unknown provenance and chronologies or single series representing potential sites of origin. Therefore, local reference chronologies form the backbone of dendro-provenancing. Reference chronologies may either be established from living trees or, for provenancing of historical and archaeological timber, they need to be constructed from series of historical or archaeological objects that are thought to represent local timber-sources. The proximity between a

reference of known and an object of unknown provenance is usually expressed in statistical terms by calculating either *t*-values or Gleichläufigkeit (percentage of common signs of year-to-year growth changes; cf. Bridge, 2012; Buras and Wilmking, 2015). The spatial distribution of matches is then visualized on maps, which allow for narrowing down the area of provenance to the best matches (Eckstein and Wrobel, 2007; Daly and Nymoen, 2008). Such dendro-provenancing relies on three assumptions:

1. Tree growth varies sufficiently within the study area, thus causing the formation of regionally or locally characteristic ring-width patterns.
2. The (dis-)similarity of tree growth can be quantified by statistical measures of proximity.
3. Highest statistical proximity indicates closest geographical neighbors.

* Correspondence to: Universität Zürich, Institut für Archäologie, Fachbereich Prähistorische Archäologie, Karl Schmid-Strasse 4, 8006 Zürich, Switzerland.
E-mail address: u.gut@posteo.de.

¹ CSR_{open} is a figure for the classification success rate of a k-NN and is described in detail in the Methods section.

A review on dendro-provenancing was provided by Bridge (2012). Recently, the ‘classical’ method described above has been refined. For example, Jansma et al. (2014) combined topographical, geomorphological and soil type information to reconstruct potential sites of origin for *Quercus petraea* and *Quercus robur* in the Roman epoch. Similarly, Eissing and Dittmar (2011) used maps of the potential natural vegetation of Thuringia and Bavaria to determine areas for medieval harvesting of *Picea abies* and *Abies alba*. In addition, they consulted historical sources on timber rafting and considered the hydrological system upstream of a city to locate likely sites of timber supply (Eissing and Dittmar, 2011). However, the original assumptions stated above remain crucial for both studies (Eissing and Dittmar, 2011; Jansma et al., 2014). For a very recent approach of assessing the reliability of best statistical matches, see Drake (2018), who developed a probabilistic framework for dendro-provenancing based on statistical hypothesis testing and Bayesian inference.

While the quantification of similarity in tree growth using proximity measures constitutes a virtually unchallenged assumption, the first and third assumptions are discussed controversially (Bridge, 2000; Haneca et al., 2005; Savva et al., 2006; Garcia-Gonzalez, 2008; Eissing, 2007; Eissing and Dittmar, 2011; Jansma et al., 2014). Concerning the first assumption of regional growth patterns, cluster analysis has sometimes been used to investigate statistical proximity in relation to the spatial distribution of site conditions, albeit often with limited success (Bridge, 2000; Haneca et al., 2005; Savva et al., 2006; Garcia-Gonzalez, 2008). Although Garcia-Gonzalez (2008) found that the clusters actually represent specific ecological conditions, in the studies by Bridge (2000) and Haneca et al. (2005) the spatially diffuse results of the cluster analysis could not be correlated with site factors. Savva et al. (2006) found that the clearest clusters followed elevational belts. Moreover, motivated by studies that showed distinct elevation-specific climate-growth relationships, models predicting the elevational provenance of ring-width series were formulated (Wilson and Hopfmüller, 2001; Frank and Esper, 2005; Eissing and Dittmar, 2011; Dittmar et al., 2012; King et al., 2013; Kolář, Čermák et al., 2017; Lyu et al., 2017). Besides elevation-specific differences, results of pointer year studies suggest distinct regional to site-specific differences in tree growth (Dittmar and Elling, 1999; Rolland et al., 2000; Neuwirth et al., 2004, 2007).

The third assumption was discussed even more controversially. Eissing and Dittmar (2011) pointed out that imported timber was frequently used for the construction of the larger, communal buildings in many old town centers of the medieval and modern epoch. Hence, in such cases reference chronologies for dendro-provenancing cannot be thought of as reflecting local tree growth (Eissing, 2007; Eissing and Dittmar, 2011; Jansma et al., 2014). Best matches to such references must be scrutinized and interpreted critically. Also, a close and strong relation of geographical and statistical proximity is not granted from an ecological point of view (Haneca et al., 2005; Bridge, 2012; Fowler and Bridge, 2015). For example, Bridge (2000) studied oak stands in eastern England and compared oak ring-width series within the studied sites as well as with other British sites, suggesting that series from Hockley Woods (Essex) featured higher statistical proximity when compared to a 270-km distant site chronology from Peckforton (Cheshire), than when compared to much closer sites inside a 100-km radius around Hockley Woods. The contradiction was explained by the fact that both Hockley Woods and Peckforton were located on well-drained, steep slopes, i.e. they were ecologically quite similar in spite of a large geographical distance. Most previous dendro-provenancing studies were conducted in areas that are under an Atlantic climate regime and have only low topographical complexity, such as the Polish, Belgian and Baltic coastal areas and their hinterland (Bridge, 2012). There are only a few such studies in Alpine and pre-Alpine environments (Eissing and Dittmar, 2011). In these latter environments, the complex topography may be advantageous for dendro-provenancing. However, even when there is high diversity of site conditions and micro climate, statistical neighbors may be located at distant sites, where similar environmental factors

limit tree growth (Bridge, 2000; Haneca et al., 2005; Boschetti-Maradi and Kontic, 2012).

In spite of all controversies, dendro-provenancing remains popular and is widely applied. Its core method (i.e., the provenancing via pairwise statistical comparisons), however, has never been evaluated in a small-scale mountain environment. Thus, the objectives of this study are:

1. To develop a statistical procedure for validating the three fundamental dendro-provenancing assumptions.
2. To investigate the suitability of different proximity measures for dendro-provenancing.
3. To assess the most problematic provenancing errors encountered to be able to determine the growth signal that is relevant for dendro-provenancing and thus provide a basis for future research dedicated to a better understanding of this signal in Alpine and pre-Alpine environments.

2. Materials and methods

2.1. Tree-ring data

779 increment cores of Norway spruce (*P. abies* (L.) H. Karst.) were collected between winter 2015 and fall 2016. At 15 sites in the foothills and mountains of the north-eastern Swiss Alps 15–32 healthy spruce with a diameter at breast height (DBH) ≥ 30 cm were sampled (Table 1, Figs. 1, A1). The sampling plots were relatively small, ranging between 1110 m² and 7840 m² and were chosen to represent typical site conditions of the area. For most sites, two cores were extracted per tree. In Sihlwald-Streuboden (sw), however, a pilot study was conducted where only one core per tree was acquired.

The measurement of total ring width and cross-dating were performed using standard dendrochronological procedures (Cook and Kairiukstis, 1990; Speer, 2010). Only a few ring-width series had to be excluded from further analysis, after they had failed final cross-dating checks with the program COFECHA (Grissino-Mayer, 2001). Finally, for trees with more than one series of radial measurements, mean ring-width series were calculated. Thus, a data set of 401 ring-width series was created (Table 2, ring-width data in the Online Supplementary Material).

2.2. Dendro-provenancing as a k-Nearest Neighbor Analysis (k-NN)

Dendro-provenancing can be framed as a *Nearest Neighbor (NN) Analysis* (Cover and Hart, 1967; Schmitt, 2006). In One Nearest Neighbor Analysis (1-NN), for example, the object being classified is assigned the same class as its statistical NN. However, in order to evaluate the robustness of a classification, NN Analysis may relate to more than just the 1-NN and inquire the consistency of classifications based on any number k of NN. If $k = 10$, for example, the 10 NN are considered for classification. Consequently, the object (ring-width series) is assigned the same class (site label) as the most abundant class among the 10 NN. This *k-Nearest Neighbor* (k-NN) approach is pursued here. The procedure involves the following steps:

1. A ring-width series is ‘anonymized’. All other series of the data set are assigned classes corresponding to their site provenance.
2. A proximity measure is chosen and values are calculated for each pairwise comparison between the anonymized series and all other series of known provenance.
3. Proximity values are sorted in ascending order and ranked accordingly. The number of highest ranking values (k) is chosen and the classes of the respective NN are evaluated. The most abundant class among those k NN determines the class attributed to the anonymized series. Thus, for each possible k (here, $k \in \{1, \dots, 400\}$), the class of the anonymous series is predicted.

Table 1

Site factors. Elevation in m a.s.l. Slope in degrees. Determination of root penetration depth (root pen. depth), nutrient storage capacity (nutr.stor.), and waterlogging (water log.), according to digital soil suitability map of Switzerland (Federal Office for Agriculture: map.geo.admin.ch, 17.8.2017). Duration of growing season in days (grow.s.), annual temperature average in °C (t.av.), and annual precipitation sums in mm (precip.), derived by calculating the mean for the period 1930–2010. Growing season calculated according to the ETCCDI definition (etccdi.pacificclimate.org, 7.9.2018) with R Package *climdex.pcic* (Bronaug, 2018). Monthly and daily mean temperatures and precipitation sums for 1930–2010 were provided by the Land Use Dynamics Research Group at WSL. These data had been derived by a spatial interpolation of data from the MeteoSwiss network using DAYMET (Thornton et al., 1997) to a grid with cell size of 1 h.

Site	Region	Elevation	Slope	Exp.	Root pen.depth	Nutr.stor.	Water log.	Grow.s.	t.av.	Precip.
hw: Glarus-Haltenwald	Linth	627	18	NE	Medium	Medium	Moist	236	8.7	1522
nb: Sool-Nuebaennli	Linth	845	40	S	Very superficial	Very Low	No moist	224	7.6	1603
how: Sool-Hohwald	Linth	1022	41	S	Very superficial	Very low	No moist	213	6.7	1712
gand: Elm-Gandwald	Linth	1180	27	E	Superficial	Low	No moist	212	6.6	1518
ww: Spiringen-Waengiwald	Linth	1707	19	N	Very superficial	Very low	No moist	144	2.4	2019
rw: Elm-Raminerwald	Linth	1723	26	SW	Superficial	Low	No moist	146	2.5	1924
bw: Schmerikon-Bannwald	Obersee	472	15	N	Deep	Medium	Low wet	242	9	1487
ew: Eschenbach-Eggwald	Obersee	618	14	NW	Medium	Low	Moist	234	8.4	1546
sb: Gommiswald-Steibruch	Obersee	856	10	NW	Deep	Medium	Low wet	213	6.8	1853
chw: Eschenbach-Cholwald	Obersee	1106	20	N	Medium	Good	Low wet	192	5.5	1853
sw: Sihlwald-Streuboden	Sihl	646	6	NE	Deep	Good	Low wet	233	8.3	1357
fri: Feusisberg-Friesischwand	Sihl	829	15	N	Medium	Medium	Low wet	219	7.2	1660
kar: Unteriberg-Karenstockwald	Sihl	1000	26	NW	Superficial	Low	Moist	193	5.5	2049
gw: Oberiberg-Gschwaendwald	Sihl	1198	13	SE	Superficial	Medium	Wet	194	5.5	2043
furg: Alphal-Furggelenstock	Sihl	1506	16	E	Superficial	Medium	Wet	171	4	2243

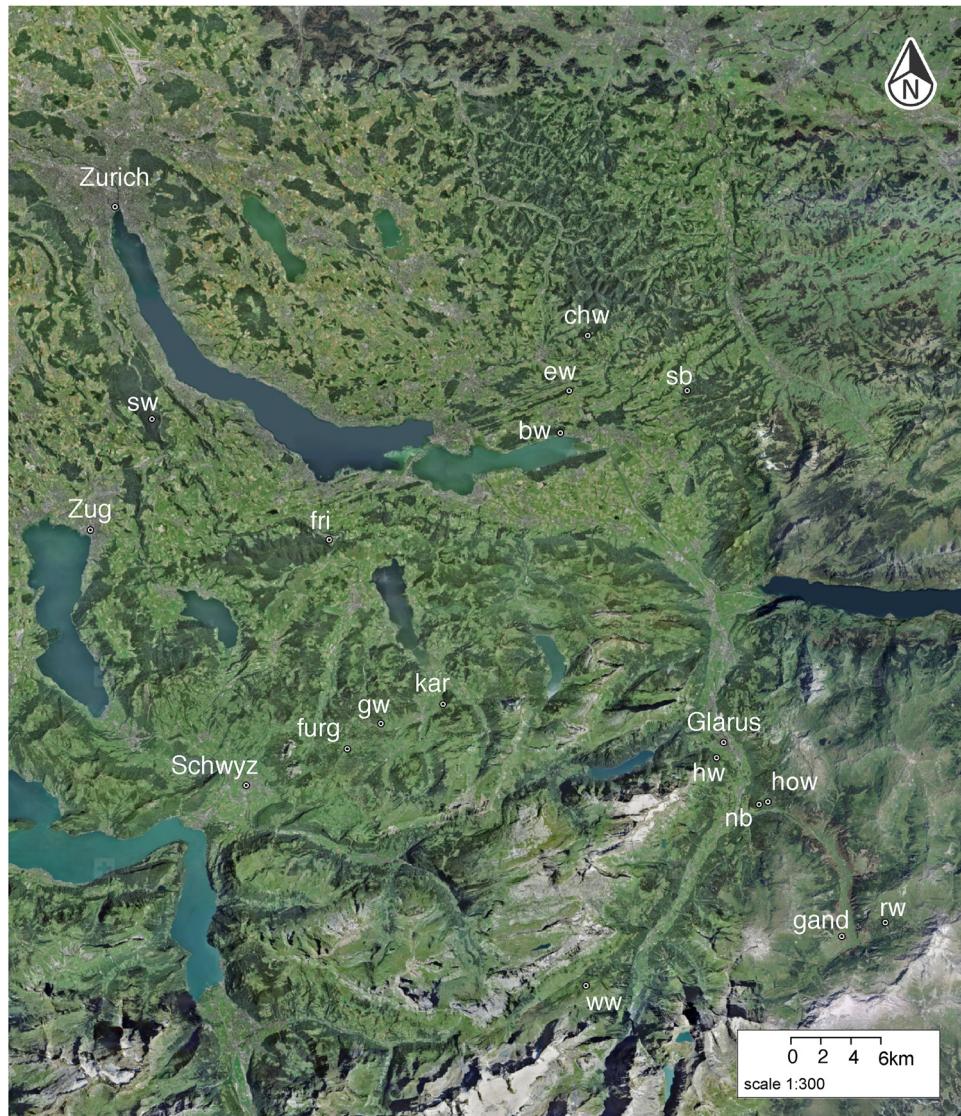


Fig. 1. Map of sites ($46^{\circ} 48' 37''$ – $47^{\circ} 29' 24''$ N, $8^{\circ} 26' 50''$ – $9^{\circ} 16' 23''$ E). Full names of sites see Table 1. Reproduced by permission of swisstopo (BA17135).

Table 2

Chronology and signal strength statistics as well as site-wise mean classification success rate (*CSR*) for k -NN stages $k = 40$ till $k = 1$ conducted with the measure tHO (Table 3) on the complete data set. Mean replication was calculated excluding the years of a chronology with a replication < 2 .

Site	<i>n</i>	Mean s.length	First yr.	Last yr.	Mean replic.	Rbar	EPS	Mean CSR
bw	27	126.33	1885	2016	25.84	0.32	0.92	87.69
chw	24	93.25	1860	2016	14.60	0.37	0.90	29.06
ew	26	105.85	1875	2016	20.19	0.36	0.92	75.38
fri	31	115.00	1880	2016	26.02	0.47	0.96	93.15
furg	19	210.63	1645	2015	11.23	0.42	0.89	81.18
gand	25	104.84	1890	2015	20.96	0.43	0.94	89.10
gw	15	127.00	1795	2015	13.38	0.38	0.89	48.33
how	23	107.78	1897	2016	20.82	0.49	0.95	62.61
hw	30	130.23	1878	2014	28.72	0.39	0.95	98.67
kar	30	81.90	1881	2016	21.36	0.39	0.93	59.17
nb	30	155.30	1851	2016	28.07	0.49	0.96	97.42
rw	32	160.16	1773	2015	21.09	0.48	0.95	95.94
sb	30	66.77	1938	2016	25.99	0.34	0.93	85.92
sw	30	97.40	1904	2014	27.27	0.26	0.90	71.92
ww	29	184.69	1755	2015	20.67	0.42	0.94	88.19

4. The classification is checked by comparing the true class (i.e., the site of provenance) of the anonymized series with its predicted class.

The k-NN is based on individual ring-width series. Reference chronologies are not used because such chronologies implicitly assume that all individual series share a site-specific growth signal, which is one of the key assumptions (c.f. Introduction) that is investigated here.

2.3. Measures of proximity

In cluster analysis, the term *proximity measure* is used to denote measures of distance, similarity and dissimilarity (see Everitt et al., 2011, p. 43). Similarly, in this paper the term *proximity measure* is used to refer to any value calculated to express statistical proximity, be it measures of distance, similarity or dissimilarity.

Most of the measures applied in this study (Table 3) are well known and frequently used in dendrochronology. Thus, the reader is relegated to the cited references and Supplementary Online Material for more

information.

Besides the measures that are based on the agreement of the growth-sign (GL and FGGL, Table 3), two correlation coefficients were investigated: Pearson's correlation coefficient (Pearson, 1895) and Spearman's rank correlation (Hollander and Wolfe, 1973; Best and Roberts, 1975; Table 3). For both correlation coefficients, *t*-values can be calculated (Kendall and Gibbons, 1990). Thus, the dependence on sample size, i.e. years of overlap, of the correlation coefficient is reduced (Wigley et al., 1987). No *p*-values are calculated here because classic dendro-provenancing has almost always been done based on *t*-values (Wazny, 2002; Haneca et al., 2005; Eckstein and Wrobel, 2007; Daly and Nymoen, 2008; Eissing and Dittmar, 2011; Hellmann et al., 2017).

Both correlation coefficients are sensitive to trends and other low frequency fluctuations (Wigley et al., 1987), thus, preprocessing is necessary. However, growth patterns of approximately 5 to 20 years in length, i.e., medium-frequency patterns, are possibly relevant for provenancing, as they may reflect similar local forest management or other growth reactions due to spatially limited disturbances, e.g. storm damages. Thus, several k-NNs were calculated from the raw ring-width series. Additionally, low-frequency tolerant preprocessing methods, such as DET and SPL67pct (Table 3), were applied. All calculations and statistical analyses were done using the statistical software R version 3.4.3 (R Core Team, 2017). The R code for the main analyses is provided in the Online Supplementary Material.

2.4. Cross-validation using scissor plots

2.4.1. The scissor plot

Because the true class (i.e., the site of provenance) of the anonymized series is known, classifications accomplished by k-NN can be validated rigorously by iteratively anonymizing each ring-width series of the data set once, then predicting its class (i.e., site) for different numbers of *k*. This kind of cross-validation method is known as *leave-one-out cross-validation* (Arlot and Celisse, 2010). From the cross-validated classifications, the *classification error rate* (CER) is derived by dividing the number of wrong classifications by the total number of classification trials.

The probability for a correct classification rises if more NN come from the same site as the anonymized ring-width series. Thus, these *on-*

Table 3

Overview of proximity measures.

Abbreviation	Correlation	<i>t</i> statistic	Preprocessing/transformation	Reference
r	Pearson	-none-	-none-	Pearson (1895)
t	Pearson	Yes	-none-	Edgett and Noon (1984)
tHO	Pearson	Yes	Log first differences	Hollstein (1980)
tDIFF	Pearson	Yes	First differences	Stock and Watson (2015)
tBP	Pearson	Yes	5 year running mean	Baillie and Pilcher (1973)
tAR	Pearson	Yes	ar-model	Brockwell and Davis (1996)
tsPL67pct	Pearson	Yes	spline 2/3 series, 0.5 freq. resp.	Cook (1981)
tsPL30yrs	Pearson	Yes	spline 30 years, 0.5 freq. resp.	Cook (1981)
tsPL10yrs	Pearson	Yes	spline 10 years, 0.5 freq. resp.	Cook (1981)
tDET	Pearson	Yes	Deterministic detrending	Fritts (1976)
tARS	Pearson	Yes	Double-detrending & ar-model	Cook (1985)
s	Spearman	-none-	-none-	Best and Roberts (1975)
ts	Spearman	Yes	-none-	Kendall and Gibbons (1990)
tsHO	Spearman	Yes	Log first differences	Hollstein (1980)
tsDIFF	Spearman	Yes	First differences	Stock and Watson (2015)
tsBP	Spearman	Yes	5 year running mean	Baillie and Pilcher (1973)
tsAR	Spearman	Yes	ar-model	Brockwell and Davis (1996)
tsPL67pct	Spearman	Yes	spline 2/3 series, 0.5 freq. resp.	Cook (1981)
tsPL30yrs	Spearman	Yes	spline 30 years, 0.5 freq. resp.	Cook (1981)
tsPL10yrs	Spearman	Yes	spline 10 years, 0.5 freq. resp.	Cook (1981)
tsDET	Spearman	Yes	Deterministic detrending	Fritts (1976)
tsARS	Spearman	Yes	Double-detrending & ar-model	Cook (1985)
GL	-none-	-none-	Falling (-) or ascending (+) yearly intervals	Huber (1943)
FGGL	-none-	-none-	Like GL but class 1 intervals (Table A1) excluded	von Jazewitsch (1948)

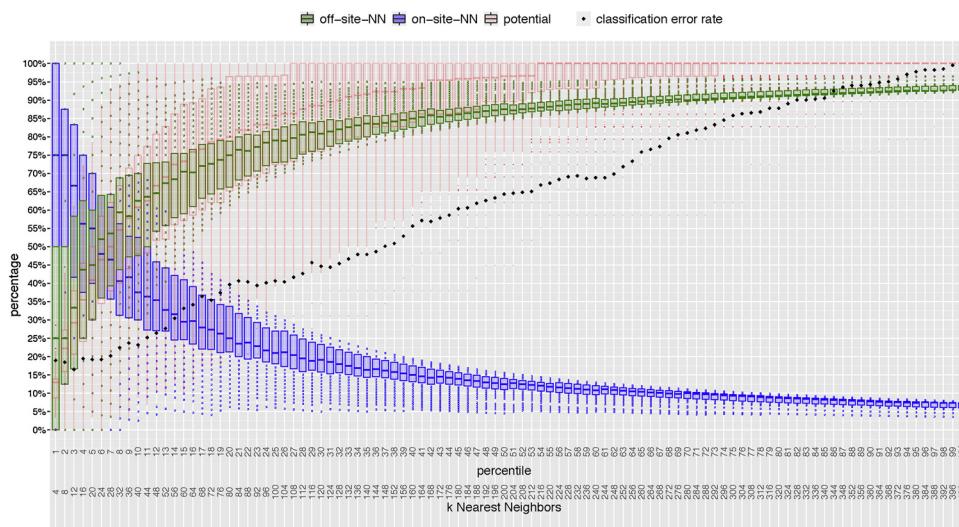


Fig. 2. Scissor plot for measure tHO; complete data set. Scissor plots visualize the leave-one-out cross-validation process. During this process, each ring-width series of the data set is anonymized once and values for on-site-NN/off-site-NN ratio and potential ratio are calculated (ratios plotted as percentages on y-axis, see Methods for details). As the complete data set encompasses $n = 401$ ring-widths series, the box-plots represent 401 values. Of the total 400 box-plots that would result from each possible setting for k , only 100 box-plots are visualized on the x-axis as the figure would get too large otherwise. This subset is called k % -NN (see Methods for details). There is no box-plot for the classification error rate as this rate is a single number, which equals the percentage of wrongly classified series at each setting for k .

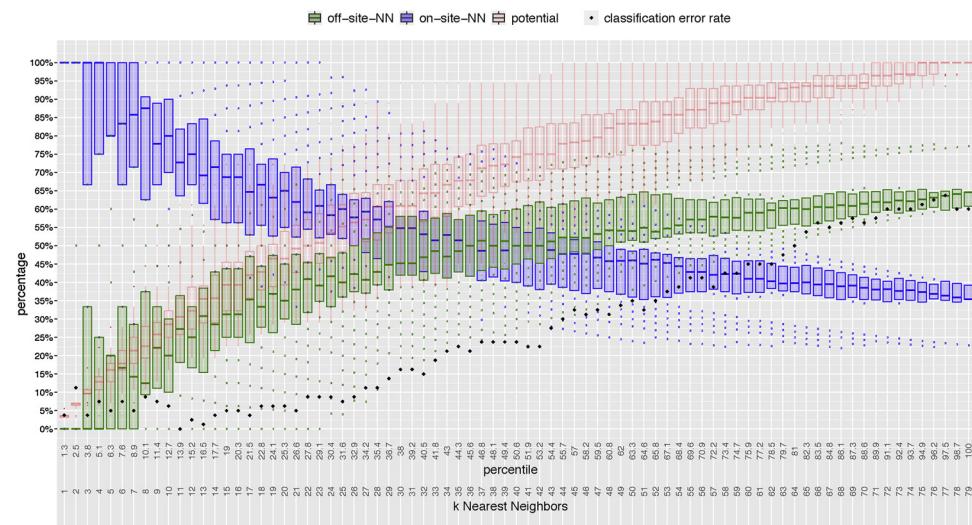


Fig. 3. Scissor plot for measure tHO; high-elevation subset.

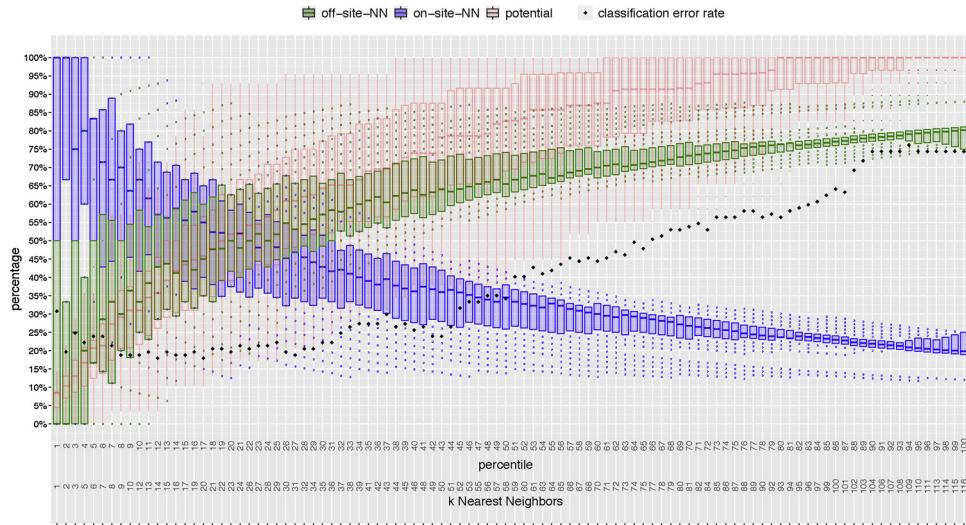


Fig. 4. Scissor plot for measure tHO; medium-elevation subset.

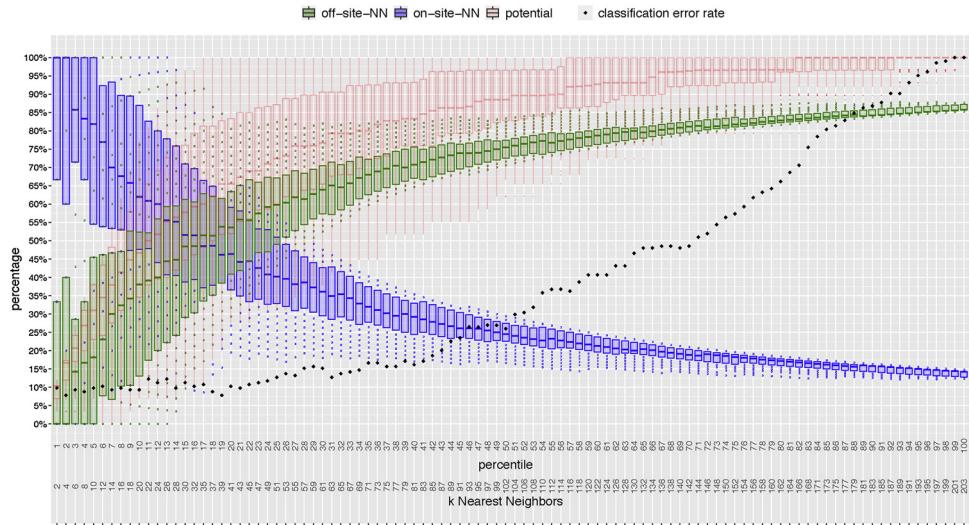


Fig. 5. Scissor plot for measure tHO; low-elevation subset.

site Nearest Neighbors (on-site-NN) are traced for all numbers of k and calculated for all iterations of the leave-one-out cross-validation (i.e., 400 iterations for the complete data set). Box plots are particularly useful to visualize the spread of the percentages of on-site-NN over all the iterations of the leave-one-out cross-validation (Fig. 2). Analogously, the number of NN coming from a different site than the anonymized ring-width series can be visualized. These NN are called *off-site Nearest Neighbors* (off-site-NN). Because the two box plot series of off-site-NN and on-site-NN resemble the crossing blades of a scissor, the result is called a *scissor plot* (Figs. 2–5). The third box plot series drawn for the *potential* is explained further below.

A complete k -NN encompasses $k_{\max} = n - 1$ classifications, where n is the sample size. As n gets larger, the classifications become increasingly challenging to visualize. Hence, often only a selection is plotted. To create a subset for k -NNs with $k_{\max} \geq 100$, the ranked proximity values are divided into percentiles and classifications are calculated for each percentile of NN. In this paper, such a subset of a k -NN is referred to as a $k\%$ -NN.

2.4.2. Rating a proximity measure with a scissor plot

The $k\%$ -NN data visualized on scissor plots (for data sets where $k_{\max} \geq 100$) represents an arbitrary selection of the data underlying a complete k -NN. However, the information of a scissor plot can be expressed in numbers, and these can be calculated from the complete k -NN data without having to plot the data.

- The rating procedure is based on four indicators (Fig. 6):
 - opening_{ratio}
 - C_{SR}
 - on – site – NN_{ratio}
 - potential_{ratio}
- These indicators are summarized by the *rating score*. In addition, C_{SRopen} provides a more intuitive rating figure (Fig. 6).

1. *opening_{ratio}*. An *opening* is defined as a classification stage $k \in \{1 \dots k_{\max}\}$ for which the 25% quantile of the on-site-NN box plot series does not overlap with the 75% quantile of the off-site-NN box plot series (for a barely opened scissor plot, cf. $k = 8$, Fig. 2; for a clear opening, cf. $k = 16$, Fig. 5). In these stages, more than 75% of all iterations of the k -NN allow for more on-site-NN than off-site-NN among the set of classifiers. Thus, the probability for a correct classification is high. The *opening_{ratio}* is the sum of open stages (N_{open}) expressed as a percentage of the total number of classification stages (k_{\max}):

$$\text{opening ratio} = \frac{N_{\text{open}}}{k_{\max}} * 100 \quad (1)$$

2. *C_{SR}*. The stability of the classification error rate (CER) or of its complement, the classification success rate (CSR), is crucial for assessing the classification performance of a proximity measure. The overall classification performance is evaluated by calculating the *mean classification success rate* (C_{SR}):

$$\bar{\text{CSR}} = \frac{\sum_{i=1}^{k_{\max}} (100 - \text{CER}_i)}{k_{\max}} \quad (2)$$

3. *on – site – NN_{ratio}*. In general, the higher the on – site – NN_{ratio}, the higher the probability of a correct classification of an anonymized ring-width series. The on-site-NN box plot series of a scissor plot shows the spread of the on-site-NN for each stage (Fig. 2). For the rating of the on-site-NN over all stages, the following ratio is calculated, where $\overline{\text{on – site – NN}_i}$ is the median on-site-NN percent at the k -NN stage $i = k \in \{1 \dots k_{\max}\}$:

$$\text{on – site – NN}_{\text{ratio}} = \frac{\sum_{i=1}^{k_{\max}} \overline{\text{on – site – NN}_i}}{k_{\max}} \quad (3)$$

4. *potential_{ratio}*. The *potential* characterizes the portion of on-site-NN at a specific stage k of the k -NN (red box plot series in Fig. 2). It takes the value of 100% if all on-site-NN in the data set are found. A high value for *potential* is a sign of a strong common growth signal within the on-site ring-width series. It is also a sign of high signal difference between on-site-NN and off-site-NN. The *potential_{ratio}* is defined as follows, where $\overline{\text{potential}_i}$ is the median potential at k -NN stage $i = k \in \{1 \dots k_{\max}\}$:

$$\text{potential}_{\text{ratio}} = \frac{\sum_{i=1}^{k_{\max}} \overline{\text{potential}_i}}{k_{\max}} \quad (4)$$

Rating score. These four indicators range from a minimum of 0% to a maximum of 100%. Thus, the overall rating score is calculated as:

$$\text{rating score} = \frac{\text{opening}_{\text{ratio}} + \bar{\text{CSR}} + \text{on – site – NN}_{\text{ratio}} + \text{potential}_{\text{ratio}}}{4} \quad (5)$$

C_{SRopen}. Although the rating score provides the mean of the key indicators of a scissor plot, this mean remains difficult to interpret. The CSR presents a much more intuitive figure. Whenever an opening is present in a scissor plot, the *mean classification success rate from the first opening up to stage k = 1* (C_{SRopen}) can be calculated. However, for some k -NN, C_{SRopen} cannot be calculated as there is no opening in the scissor

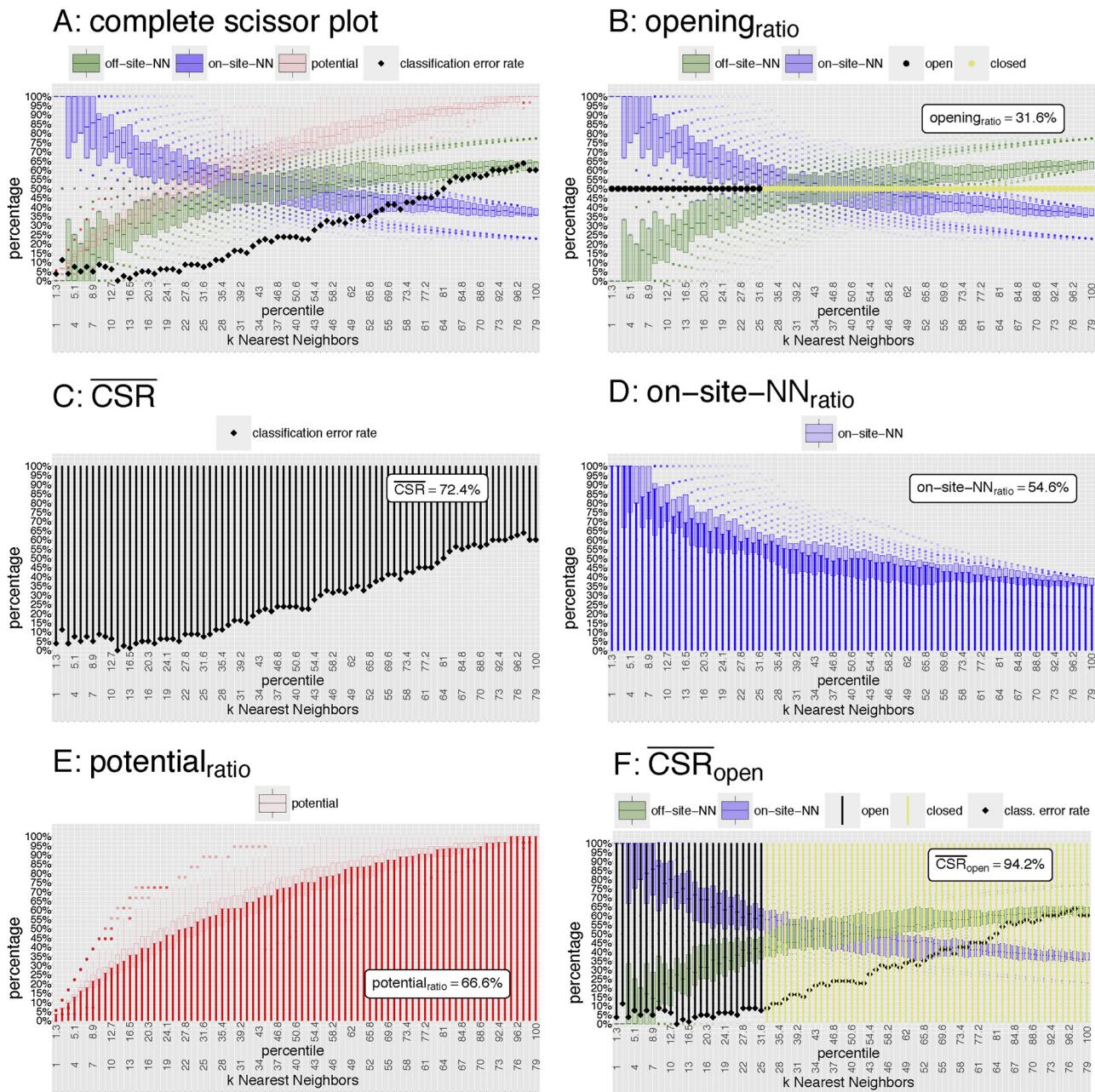


Fig. 6. Indicators used for rating a scissor plot (see Methods for detailed mathematical definitions). Subplot A shows a complete scissor plot (i.e., the plot for tHO; high-elevation subset) with all indicators on one plot. Subplots B–F illustrate how, for each indicator, the numeric score is connected to the visual representation on the scissor plot.

plot.

2.4.3. Applying k-NN to the data set

In a first step, k-NNs were calculated on the complete data set for every proximity measure (Table 3). In a second step, k-NNs were calculated on regional subsets. The three subsets, i.e. Sihl, Linth and Obersee, represent three different watersheds in the study area (for assignment of site to region cf. Table 1, Fig. 1). Because tree growth strongly depends on elevation (Frank and Esper, 2005; Savva et al., 2006; Dittmar et al., 2012; King et al., 2013; Kolář, Čermák et al., 2017; Lyu et al., 2017), in a third step k-NNs were calculated for elevation-specific subsamples of the data set. These subsamples correspond to three elevational bands, representing low (< 1000 m a.s.l.), medium (1000–1500 m a.s.l.) and high elevation sites (> 1500 m a.s.l.). To identify the site pairs that were most difficult for the classification, of

each site 15 series were randomly selected and scissor plots were drawn for all possible site pairs. Then, the rating scores were visualized via a so-called *heat plot* (see Section 3.4).

Due to space constraints, scissor plots are shown for tHO only (Figs. 2–5). For the rest of the proximity measures, the results are presented based on the rating scores (Tables 4–6).

3. Results

Firstly, the results for the complete data set are presented. They provide the basis for discussing the suitability of different proximity measures for k-NN classification (see Section 4). Secondly, the results for the regional subsamples, and thirdly, the results for the elevational subsamples are presented. The two subsampling strategies aim at detecting regional and elevational differences in the k-NN classification

Table 4

Rating summary part 1. For information on proximity measures see Table 3. For details on the abbreviated rating figures see Method section.

All sites	r	t	s	ts	tBP	tsBP	tHO	tsHO	tDIFF	tsDIFF	tAR	tsAR
	-none-	-none-	-none-	-none-	0.2	0.2	0.8	0.8	0.8	0.8	0.8	0.8
opening _{ratio}	-none-	-none-	-none-	-none-	0.2	0.2	0.8	0.8	0.8	0.8	0.8	0.8
on – site – NN _{ratio}	12.9	13.1	12.6	13.1	17.9	18.3	18.7	18.7	18.6	18.7	18.5	18.6
potential _{ratio}	72.8	73.1	72.8	73.3	86.7	87.4	88.2	88.1	87.6	88.2	87.5	88.2
C _{SR}	24.0	24.1	23.6	24.7	36.0	36.6	38.2	38.0	39.5	39.1	40.6	41.4
rating score	27.4	27.6	27.3	27.8	35.2	35.6	36.5	36.4	36.6	36.7	36.9	37.2
C _{SR} open	-none-	-none-	-none-	-none-	76.3	75.3	79.2	76.9	75.9	76.7	78.5	76.0
All sites	tARS	tsARS	tDET	tsDET	tSPL67pct	tsSPL67pct	tSPL30yrs	tsSPL30yrs	tSPL10yrs	tsSPL10yrs	GL	FGGL
	0.5	0.5	-none-	-none-	-none-	-none-	-none-	-none-	0.8	0.5	-none-	-none-
opening _{ratio}	0.5	0.5	-none-	-none-	-none-	-none-	-none-	-none-	0.8	0.5	-none-	-none-
on – site – NN _{ratio}	17.9	18.1	13.5	13.5	15.4	15.4	16.7	16.6	18.4	18.4	15.4	16.5
potential _{ratio}	87.8	87.7	73.7	74.3	80.7	81.2	84.2	84.1	87.6	87.9	83.0	84.9
C _{SR}	36.3	37.1	28.5	30.1	34.0	34.5	33.8	34.2	36.7	37.5	35.5	35.6
rating score	35.6	35.8	28.9	29.5	32.5	32.8	33.7	33.7	35.9	36.1	33.5	34.2
C _{SR} open	75.3	72.0	-none-	-none-	-none-	-none-	-none-	-none-	75.7	71.8	-none-	-none-
Sihl subset	r	t	s	ts	tBP	tsBP	tHO	tsHO	tDIFF	tsDIFF	tAR	tsAR
	-none-	-none-	-none-	-none-	1.6	9.7	11.3	14.5	16.9	19.4	18.5	19.4
opening _{ratio}	-none-	-none-	-none-	-none-	1.6	9.7	11.3	14.5	16.9	19.4	18.5	19.4
on – site – NN _{ratio}	37.7	38.1	37.8	38.2	43.3	43.1	44.9	44.7	45.4	45.2	44.8	44.9
potential _{ratio}	71.5	72.0	72.0	72.6	77.9	77.5	79.2	79.0	79.7	79.4	79.0	79.0
C _{SR}	47.9	48.5	47.5	48.4	55.1	54.3	57.8	57.5	61.2	59.6	60.9	60.6
rating score	39.3	39.6	39.3	40.2	46.5	46.5	49.1	49.5	51.4	50.1	50.8	51.0
C _{SR} open	-none-	-none-	-none-	-none-	70.1	85.2	84.9	88.4	90.0	88.7	89.7	91.4
Sihl subset	tARS	tsARS	tDET	tsDET	tSPL67pct	tsSPL67pct	tSPL30yrs	tsSPL30yrs	tSPL10yrs	tsSPL10yrs	GL	FGGL
	14.5	13.7	0.8	0.8	4.0	4.0	4.0	2.4	10.5	10.5	2.4	-none-
opening _{ratio}	14.5	13.7	0.8	0.8	4.0	4.0	4.0	2.4	10.5	10.5	2.4	-none-
on – site – NN _{ratio}	43.3	43.6	37.4	38.1	38.3	39.1	38.9	39.3	43.1	43.1	39.5	41.2
potential _{ratio}	77.7	78.2	71.3	72.1	71.4	72.2	72.1	72.7	76.9	77.2	73.5	75.7
C _{SR}	57.3	58.4	47.1	47.3	53.5	55.3	55.5	55.5	55.5	56.1	51.2	52.8
rating score	48.2	48.5	39.2	39.6	41.8	42.6	42.6	42.5	46.5	46.7	41.6	42.4
C _{SR} open	87.4	87.3	68.5	75.2	78.9	75.8	79.2	76.0	85.0	85.8	73.7	-none-
Linth subset	r	t	s	ts	tBP	tsBP	tHO	tsHO	tDIFF	tsDIFF	tAR	tsAR
	-none-	0.6	-none-	-none-	17.3	19.6	16.7	18.5	16.1	16.1	17.3	17.9
opening _{ratio}	-none-	0.6	-none-	-none-	17.3	19.6	16.7	18.5	16.1	16.1	17.3	17.9
on – site – NN _{ratio}	30.8	31.0	30.0	30.3	40.9	41.2	41.3	41.1	40.7	41.1	40.8	41.3
potential _{ratio}	74.7	75.0	73.2	73.7	86.1	86.3	86.4	86.6	86.1	86.5	85.8	86.5
C _{SR}	44.2	44.7	43.8	45.5	57.6	59.7	56.3	57.5	54.0	57.4	57.5	59.0
rating score	37.4	37.8	36.8	37.4	50.4	51.7	50.2	50.9	49.2	50.2	50.4	51.2
C _{SR} open	-none-	77.5	-none-	-none-	89.2	89.8	90.8	89.5	87.8	88.5	91.6	90.2
Linth subset	tARS	tsARS	tDET	tsDET	tSPL67pct	tsSPL67pct	tSPL30yrs	tsSPL30yrs	tSPL10yrs	tsSPL10yrs	GL	FGGL
	14.9	15.5	1.2	1.2	5.4	3.0	10.7	10.1	16.1	20.2	1.8	13.7
opening _{ratio}	14.9	15.5	1.2	1.2	5.4	3.0	10.7	10.1	16.1	20.2 <td>1.8</td> <td>13.7</td>	1.8	13.7
on – site – NN _{ratio}	41.2	41.4	32.3	32.4	37.4	37.7	39.4	39.7	41.3	41.5	36.7	38.6
potential _{ratio}	86.1	86.7	75.5	75.7	81.9	81.8	84.6	84.7	86.5	86.8	83.0	84.5
C _{SR}	55.0	57.0	49.1	50.5	55.0	55.6	55.0	55.5	56.7	58.2	57.6	58.8
rating score	49.3	50.1	39.5	39.9	44.9	44.5	47.4	47.5	50.1	51.7	44.8	48.9
C _{SR} open	89.8	90.5	72.2	70.2	82.5	81.8	86.8	86.5	89.4	90.2	73.5	87.0
Obersee subset	r	t	s	ts	tBP	tsBP	tHO	tsHO	tDIFF	tsDIFF	tAR	tsAR
	-none-	-none-	-none-	-none-	11.3	13.2	22.6	17.9	25.5	21.7	17.9	20.8
opening _{ratio}	-none-	-none-	-none-	-none-	11.3	13.2	22.6	17.9	25.5	21.7	17.9	20.8
on – site – NN _{ratio}	37.1	37.1	36.8	37.0	44.5	45.9	46.5	45.4	47.1	46.5	46.1	46.2
potential _{ratio}	65.1	65.1	64.9	65.1	74.0	75.2	75.6	75.0	76.0	75.9	75.4	75.4
C _{SR}	50.4	50.8	49.6	50.1	63.9	67.1	64.9	68.6	69.3	71.3	66.9	69.9
rating score	38.2	38.2	37.8	38.1	48.4	50.4	52.4	51.7	54.4	53.9	51.6	53.1
C _{SR} open	-none-	-none-	-none-	-none-	81.1	84.2	84.9	87.5	89.7	90.1	87.4	89.1
Obersee subset	tARS	tsARS	tDET	tsDET	tSPL67pct	tsSPL67pct	tSPL30yrs	tsSPL30yrs	tSPL10yrs	tsSPL10yrs	GL	FGGL
	14.2	6.6	-none-	-none-	1.9	-none-	4.7	2.8	20.8	20.8	-none-	0.9
opening _{ratio}	44.9	44.4	35.2	35.9	41.0	40.9	42.9	42.0	46.2	46.1	40.9	41.7
potential _{ratio}	74.1	74.2	62.1	63.0	69.4	69.5	70.9	71.0	75.1	75.6	71.8	71.4
C _{SR}	67.4	66.7	49.3	50.0	59.8	58.7	63.6	63.6	67.5	68.9	63.1	61.2
rating score	50.1	48.0	36.7	37.2	43.0	42.3	45.5	44.9	52.4	52.8	44.0	43.8
C _{SR} open	86.1	80.2	-none-	-none-	70.7	-none-	80.1	76.4	88.3	87.1	-none-	66.7

Table 5

Rating summary part 2. For information on proximity measures see [Table 3](#). For details on the abbreviated rating figures see Method section.

High-elevation sites	r	t	s	ts	tBP	tsBP	tHO	tsHO	tDIFF	tsDIFF	tAR	tsAR
opening _{ratio}	-none-	1.3	-none-	-none-	24.1	31.6	31.6	25.3	26.6	27.8	26.6	24.1
on – site – NN _{ratio}	43.5	44.4	43.6	43.8	54.2	56.4	54.6	54.6	53.2	54.0	54.7	55.8
potential _{ratio}	57.7	57.8	58.1	58.2	66.2	68.0	66.6	66.7	66.1	66.4	66.6	68.2
C _{SR}	49.0	49.2	49.9	50.2	69.2	72.7	72.4	71.5	69.4	70.6	71.1	71.4
rating score	37.6	38.2	37.9	38.1	53.4	57.2	56.3	54.5	53.8	54.7	54.7	54.9
C _{SR} _{open}	-none-	76.2	-none-	-none-	91.1	91.2	94.2	90.5	90.7	89.3	91.3	89.5
High-elevation sites	tARS	tsARS	tDET	tsDET	tSPL67pct	tsSPL67pct	tSPL30yrs	tsSPL30yrs	tSPL10yrs	tsSPL10yrs	GL	FGGL
opening _{ratio}	30.4	27.8	3.8	-none-	8.9	15.2	20.3	30.4	22.8	30.4	6.3	22.8
on – site – NN _{ratio}	55.6	56.4	47.0	46.4	51.1	52.4	54.8	56.8	54.9	56.2	49.6	52.3
potential _{ratio}	67.4	68.3	59.9	60.2	64.4	65.8	66.7	68.7	66.8	68.1	63.8	64.6
C _{SR}	70.8	72.0	53.9	55.6	62.5	65.7	68.1	70.5	68.6	71.9	65.4	70.2
rating score	56.0	56.2	41.1	40.5	46.7	49.8	52.4	56.6	53.3	56.6	46.3	52.5
C _{SR} _{open}	92.4	91.0	71.2	-none-	80.9	82.7	86.4	87.1	92.4	91.8	77.3	88.9
Medium-elevation sites	r	t	s	ts	tBP	tsBP	tHO	tsHO	tDIFF	tsDIFF	tAR	tsAR
opening _{ratio}	-none-	-none-	-none-	-none-	2.6	0.9	2.6	2.6	6.0	3.4	5.2	4.3
on – site – NN _{ratio}	32.5	32.7	32.6	32.8	37.8	38.0	38.8	38.8	37.5	37.9	37.6	38.1
potential _{ratio}	68.0	68.6	67.7	68.1	72.9	74.1	73.8	74.8	73.9	74.9	73.8	74.3
C _{SR}	37.3	36.6	36.7	36.1	56.7	56.6	59.5	60.2	66.0	63.4	63.4	62.2
rating score	34.5	34.5	34.3	34.2	42.5	42.4	43.7	44.1	45.9	44.9	45.0	44.7
C _{SR} _{open}	-none-	-none-	-none-	-none-	75.0	76.1	77.1	79.0	81.6	79.4	80.3	81.9
Medium-elevation sites	tARS	tsARS	tDET	tsDET	tSPL67pct	tsSPL67pct	tSPL30yrs	tsSPL30yrs	tSPL10yrs	tsSPL10yrs	GL	FGGL
opening _{ratio}	2.6	0.9	-none-	-none-	0.9	-none-	2.6	-none-	3.4	4.3	-none-	-none-
on – site – NN _{ratio}	38.0	37.1	32.7	33.4	35.0	33.7	36.5	35.5	38.7	38.4	31.4	33.7
potential _{ratio}	74.0	73.3	67.7	68.0	70.1	69.4	71.1	70.5	73.2	73.9	67.4	68.9
C _{SR}	57.1	55.7	42.5	44.5	52.5	52.1	57.3	56.6	58.5	59.4	48.6	59.1
rating score	42.9	41.7	35.7	36.5	39.6	38.8	41.9	40.7	43.5	44.0	36.9	40.4
C _{SR} _{open}	74.2	76.1	-none-	-none-	68.1	-none-	75.9	-none-	78.3	77.4	-none-	-none-
Low-elevation sites	r	t	s	ts	tBP	tsBP	tHO	tsHO	tDIFF	tsDIFF	tAR	tsAR
opening _{ratio}	-none-	-none-	-none-	-none-	1.5	3.0	6.4	5.4	6.9	7.4	5.4	5.4
on – site – NN _{ratio}	25.8	26.1	25.1	25.7	32.0	32.4	32.6	32.6	33.2	33.6	32.7	33.3
potential _{ratio}	72.2	72.5	71.0	71.8	79.2	80.6	80.7	80.7	81.3	81.9	80.5	81.7
C _{SR}	42.8	43.2	41.7	43.1	53.4	53.4	61.6	59.2	64.4	61.3	63.9	63.5
rating score	35.2	35.4	34.5	35.1	41.5	42.3	45.3	44.5	46.5	46.0	45.6	46.0
C _{SR} _{open}	-none-	-none-	-none-	-none-	79.2	82.5	89.3	87.0	89.4	88.2	87.3	85.7
Low-elevation sites	tARS	tsARS	tDET	tsDET	tSPL67pct	tsSPL67pct	tSPL30yrs	tsSPL30yrs	tSPL10yrs	tsSPL10yrs	GL	FGGL
opening _{ratio}	3.4	2.0	-none-	-none-	-none-	-none-	1.0	1.0	5.9	4.9	-none-	-none-
on – site – NN _{ratio}	31.3	31.9	23.9	24.4	27.0	27.6	28.9	29.0	32.3	32.4	27.8	29.7
potential _{ratio}	79.1	80.3	67.4	68.4	72.6	73.9	75.6	76.1	80.1	80.8	76.4	77.7
C _{SR}	57.1	56.3	44.3	46.0	50.2	49.3	48.5	48.9	56.7	56.9	53.5	54.7
rating score	42.7	42.6	33.9	34.7	37.5	37.7	38.5	38.8	43.7	43.7	39.4	40.5
C _{SR} _{open}	83.6	79.8	-none-	-none-	-none-	-none-	72.5	73.2	88.1	86.3	-none-	-none-

performance (discussed in Section 4).

Finally, the results emerging from investigating classification errors are presented. Because performance was low for the medium-elevation subsample (see Section 3.3), this subsample was further split into one consisting of the sites between 1000 and 1106 m, and the other between 1180 and 1198 m. Surprisingly, these subsamples posed no difficulties for the classification. To identify other subsets of sites that are particularly difficult for the classification, k-NNs were calculated for all site pairs. The effect of the thus identified, problematic sites on classification performance was quantified by dropping the respective sites from the complete data set and calculating k-NNs anew. Based on these results potential causes underlying the classification errors are discussed in Section 4.

3.1. Complete data set

All k-NNs calculated on the complete data showed poor

performance using correlation-based proximity measures that lacked preprocessing (see upper two blocks of [Table 4](#); i.e., r, s, t and ts). None of these measures showed an opening and the rating score was quite low (cf. [Table 4](#)). In fact, the score increased progressively the more rigorously the preprocessing removed the medium- and low-frequency variation. The highest rankings were accomplished by measures that were based on either AR, ARS, BP, SPL10yrs, HO or DIFF preprocessing whereas preprocessing methods that removed only little of the medium-to low-frequency variation such as DET or SPL67pct performed worse ([Table 4](#); [Fig. 2](#)). Particularly, they showed no opening and reached low rating scores only.

Differences in performance depending on the correlation coefficient in use were rather small (cf. [Table 4](#)). Except for the DIFF methods, all measures using the Pearson correlation coefficient attained a slightly higher C_{SR}_{open} than their Spearman counterparts. The rating score, however, was very similar for both correlation coefficients, and for both coefficients openings were limited to stages with small values of k

Table 6

Rating summary part 3. For information on proximity measures see Table 3. For details on the abbreviated rating figures see Method section.

Subset 1000–1106 m	r	t	s	ts	tBP	tsBP	tHO	tsHO	tDIFF	tsDIFF	tAR	tsAR
opening _{ratio}	1.3	2.6	3.9	2.6	26.3	39.5	27.6	36.8	35.5	38.2	40.8	38.2
on – site – NN _{ratio}	49.2	49.7	50.4	50.6	54.5	55.8	56.6	57.1	56.0	56.4	56.6	57.0
potential _{ratio}	62.5	62.7	62.6	62.7	69.7	71.9	70.5	71.3	70.5	71.9	71.1	71.2
C _{SR}	58.6	58.5	59.9	59.7	75.6	77.1	76.3	79.8	81.2	81.2	81.0	81.0
rating score	42.9	43.4	44.2	43.9	56.5	61.1	57.8	61.3	60.8	61.9	62.4	61.8
C _{SR} _{open}	71.0	72.7	72.0	73.0	87.9	88.3	88.5	92.7	94.9	93.9	92.8	92.7
Subset 1000–1106 m	tARS	tsARS	tDET	tsDET	tSPL67pct	tsSPL67pct	tSPL30yrs	tsSPL30yrs	tSPL10yrs	tsSPL10yrs	GL	FGGL
opening _{ratio}	31.6	31.6	5.3	7.9	18.4	11.8	22.4	26.3	28.9	34.2	13.2	21.1
on – site – NN _{ratio}	55.8	55.0	50.7	51.1	53.1	52.5	53.9	54.4	55.7	56.6	50.4	51.4
potential _{ratio}	69.7	70.2	65.8	66.7	66.9	67.1	67.0	67.7	69.5	70.8	66.9	68.3
C _{SR}	76.6	76.3	65.0	64.1	69.5	69.5	72.1	72.0	77.2	77.5	71.4	77.5
rating score	58.4	58.3	46.7	47.5	52.0	50.2	53.8	55.1	57.8	59.8	50.5	54.6
C _{SR} _{open}	89.2	91.1	77.7	79.7	81.6	82.1	86.4	84.1	90.6	89.2	81.1	87.0
Subset 1180–1198 m	r	t	s	ts	tBP	tsBP	tHO	tsHO	tDIFF	tsDIFF	tAR	tsAR
opening _{ratio}	69.2	69.2	69.2	69.2	46.2	33.3	59.0	61.5	69.2	69.2	69.2	69.2
on – site – NN _{ratio}	85.6	85.6	85.7	85.9	85.7	84.8	85.6	86.3	79.7	82.4	80.3	82.3
potential _{ratio}	70.3	70.3	70.1	70.1	66.9	65.9	67.3	68.1	67.0	67.5	68.0	68.0
C _{SR}	83.8	83.7	83.8	83.9	74.1	71.0	78.8	79.5	85.5	83.4	84.7	83.7
rating score	77.2	77.2	77.2	77.3	68.2	63.7	72.7	73.9	75.4	75.6	75.6	75.8
C _{SR} _{open}	93.3	93.2	93.5	93.6	83.8	79.1	89.7	89.2	95.7	92.7	94.6	93.1
Subset 1180–1198 m	tARS	tsARS	tDET	tsDET	tSPL67pct	tsSPL67pct	tSPL30yrs	tsSPL30yrs	tSPL10yrs	tsSPL10yrs	GL	FGGL
opening _{ratio}	69.2	69.2	61.5	69.2	69.2	69.2	61.5	61.5	41.0	51.3	56.4	46.2
on – site – NN _{ratio}	83.6	83.5	81.3	79.9	79.6	78.4	78.0	75.4	85.6	84.1	79.0	80.0
potential _{ratio}	68.2	67.5	68.4	68.8	66.9	66.6	66.2	64.0	66.8	66.0	65.1	65.8
C _{SR}	84.6	83.9	79.1	80.7	84.0	82.0	80.1	78.0	74.7	73.7	75.1	71.5
rating score	76.4	76.1	72.6	74.6	74.9	74.1	71.5	69.7	67.0	68.8	68.9	65.8
C _{SR} _{open}	94.4	93.4	89.9	89.3	93.5	90.6	91.1	87.7	88.1	81.5	82.8	77.5
Data without subset 1000–1106 m	r	t	s	ts	tBP	tsBP	tHO	tsHO	tDIFF	tsDIFF	tAR	tsAR
opening _{ratio}	-none-	-none-	-none-	-none-	1.9	2.8	4.3	4.3	5.0	5.0	5.0	4.0
on – site – NN _{ratio}	16.5	16.8	16.4	16.7	23.5	24.0	23.9	23.9	24.0	24.1	23.8	24.1
potential _{ratio}	73.8	74.1	73.3	74.0	88.5	89.4	89.4	89.9	89.2	90.0	88.8	89.6
C _{SR}	29.4	29.8	29.5	30.6	41.2	41.4	43.9	43.0	45.6	44.3	46.6	46.6
rating score	29.9	30.2	29.8	30.3	38.8	39.4	40.4	40.3	40.9	40.9	41.0	41.1
C _{SR} _{open}	-none-	-none-	-none-	-none-	81.1	83.6	89.3	87.7	89.3	88.7	88.5	86.5
Data without subset 1000–1106 m	tARS	tsARS	tDET	tsDET	tSPL67pct	tsSPL67pct	tSPL30yrs	tsSPL30yrs	tSPL10yrs	tsSPL10yrs	GL	FGGL
opening _{ratio}	3.4	2.8	-none-	-none-	-none-	-none-	0.9	0.9	3.7	4.0	-none-	-none-
on – site – NN _{ratio}	23.4	23.7	17.2	17.4	20.0	20.2	21.6	21.4	23.7	23.8	20.7	21.9
potential _{ratio}	89.5	89.9	74.7	75.4	82.5	83.3	86.1	86.2	89.1	89.5	85.5	87.1
C _{SR}	42.2	42.3	33.8	35.1	39.5	39.4	38.5	38.7	42.1	43.0	40.9	41.1
rating score	39.6	39.7	31.4	32.0	35.5	35.7	36.8	36.8	39.6	40.1	36.8	37.5
C _{SR} _{open}	87.2	83.4	-none-	-none-	-none-	-none-	76.4	75.6	87.5	85.9	-none-	-none-
Data without chw, gw, how, kar	r	t	s	ts	tBP	tsBP	tHO	tsHO	tDIFF	tsDIFF	tAR	tsAR
opening _{ratio}	-none-	-none-	-none-	-none-	3.2	3.9	5.2	5.2	5.8	5.5	5.2	4.5
on – site – NN _{ratio}	17.2	17.6	17.2	17.5	24.8	25.2	25.0	25.2	25.1	25.4	25.0	25.4
potential _{ratio}	72.8	73.2	72.5	73.1	88.5	89.3	89.2	89.6	88.8	89.6	88.4	89.1
CSR	31.1	31.3	31.1	32.2	43.1	43.2	45.8	44.5	46.9	45.8	48.0	48.0
rating score	30.3	30.5	30.2	30.7	39.9	40.4	41.3	41.1	41.6	41.6	41.6	41.8
C _{SR} _{open}	-none-	-none-	-none-	-none-	83.1	85.4	91.0	89.0	90.0	89.1	88.9	86.9
Data without chw, gw, how, kar	tARS	tsARS	tDET	tsDET	tSPL67pct	tsSPL67pct	tSPL30yrs	tsSPL30yrs	tSPL10yrs	tsSPL10yrs	GL	FGGL
opening _{ratio}	3.6	2.9	-none-	-none-	-none-	-none-	1.0	1.0	5.2	5.5	-none-	0.3
on – site – NN _{ratio}	24.6	24.9	18.0	18.3	21.0	21.4	22.8	22.9	24.9	25.2	21.8	23.2
potential _{ratio}	89.0	89.4	73.9	74.6	81.8	82.6	86.1	86.2	89.2	89.4	85.3	87.0
C _{SR}	43.6	43.7	35.2	36.6	41.1	40.9	39.8	40.1	43.9	44.5	42.6	43.0
rating score	40.2	40.2	31.8	32.4	36.0	36.2	37.4	37.5	40.8	41.2	37.4	38.4
C _{SR} _{open}	86.8	83.9	-none-	-none-	-none-	-none-	76.9	77.0	89.5	87.9	-none-	75.4

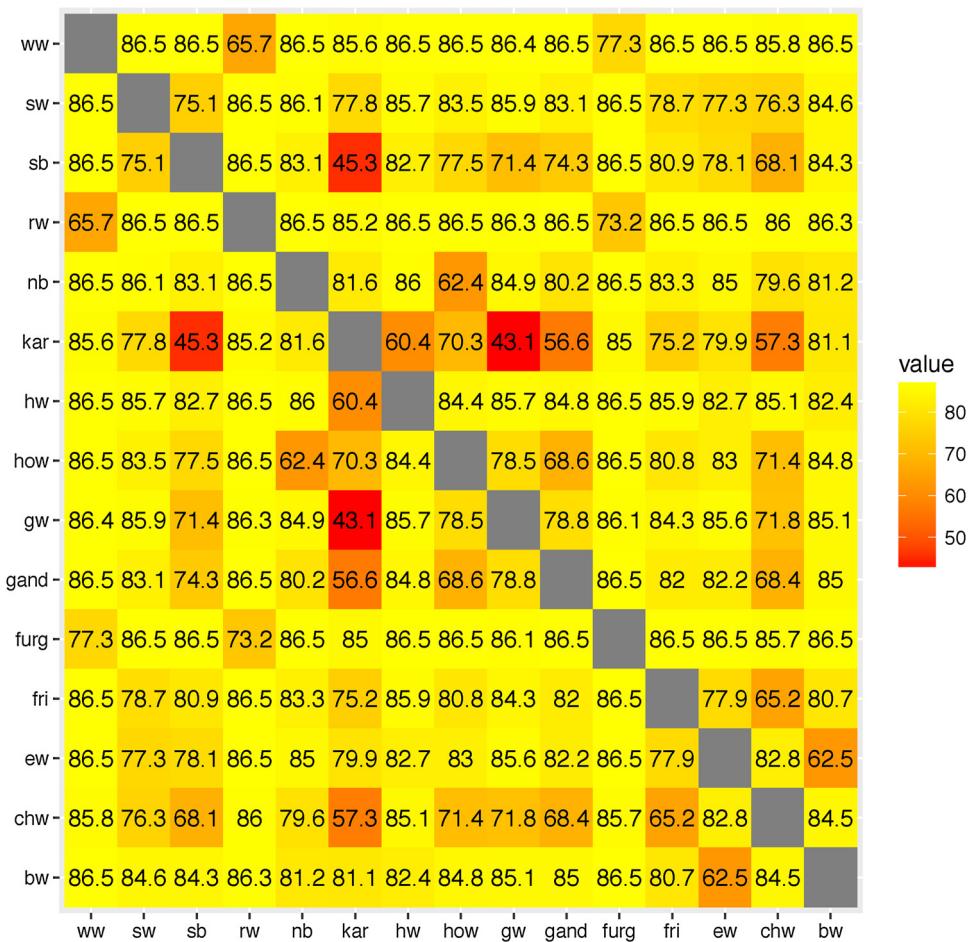


Fig. 7. Heat plot with rating scores accomplished by tHO. Rating scores were calculated from pairwise scissor plots that were drawn for all sites of the data set.

(≤ 5).

The performance of the sign-based proximity measures GL and FGGL lay somewhat above the performance of the measures that relied on DET or SPL67pct preprocessing. However, both GL and FGGL performed much better than the correlation-based methods that lacked preprocessing (i.e., r , s and t). Still, neither GL nor FGGL showed an opening. Additionally, the rating score of these two measures was distinctly lower than that of the correlation-based proximity measures using high-pass filters.

3.2. Regional subsamples

There were few differences in classification performance between the regions (see lower six blocks of Table 4). The subsample Linth had a higher k-NN performance than the other two subsamples, i.e. Sihl and Obersee. The more rigid detrending methods, e.g. SPL67pct, and the sign-based methods, especially FGGL, performed quite well with the Linth subset. For all measures, the overall rating scores were somewhat higher for the regional subsets than for the complete data set. Similar to the complete data set, correlation-based proximity measures that used high-pass filters performed best.

3.3. Elevation-specific subsamples

For the high-elevation subsample, k-NN classification was quite successful. Scissor plots opened for most of the measures (Table 5). Rating scores and CSR_{open} reached relatively high values for some measures that in other subsamples performed only poorly, such as tSPL67pct, tsSPL67pct, tDET, FGGL and GL. Again, high-pass filter

based measures performed best, with some accomplishing $\text{CSR}_{\text{open}} \geq 90\%$ (e.g. tHO, tAR, sAR, tSPL10yrs).

Within the medium-elevation subsample, k-NN classification issues were more common than within the other two elevation-specific subsamples (Table 5; cf. Figs. 3–5). The performance for the medium-elevation subsample dropped for all measures when compared to the performance reached for the high-elevation subsample. In addition, for most measures the performance was lower with the medium-elevation subsample compared to the low-elevation subsample. This was especially true for the best performing measures tHO, tsHO, tAR, tsAR, tDIFF, tsDIFF, tSPL10yrs and tsSPL10yrs.

On the low-elevation subsample, only correlation-based measures that relied on high-pass filters performed well (Table 5). Consequently, rating scores decreased rapidly the more low-frequency variation was left in an index series (SPL67pct, DET, ts, t , r , s). The sign-based measures both did not show any opening. FGGL again reached a distinctly higher rating score than GL. The rating score of FGGL even lay above that of the low-frequency tolerant, correlation-based measures (SPL67pct, DET, ts, t , r , s).

3.4. Classification errors

Heat plots of all high-performance measures (i.e., tHO, tsHO, tDIFF, tsDIFF, tAR, tsAR and tSPL10yrs) were very similar; thus only one example is shown (tHO, Fig. 7). The generally high rating scores indicated that most pairs of sites were distinguished easily via k-NN. However, there were also some lower rating scores, most of which were associated with the site Unteriberg-Karrenstock (*kar*). A second, less pronounced set of low rating scores was found for Eschenbach-Cholwald

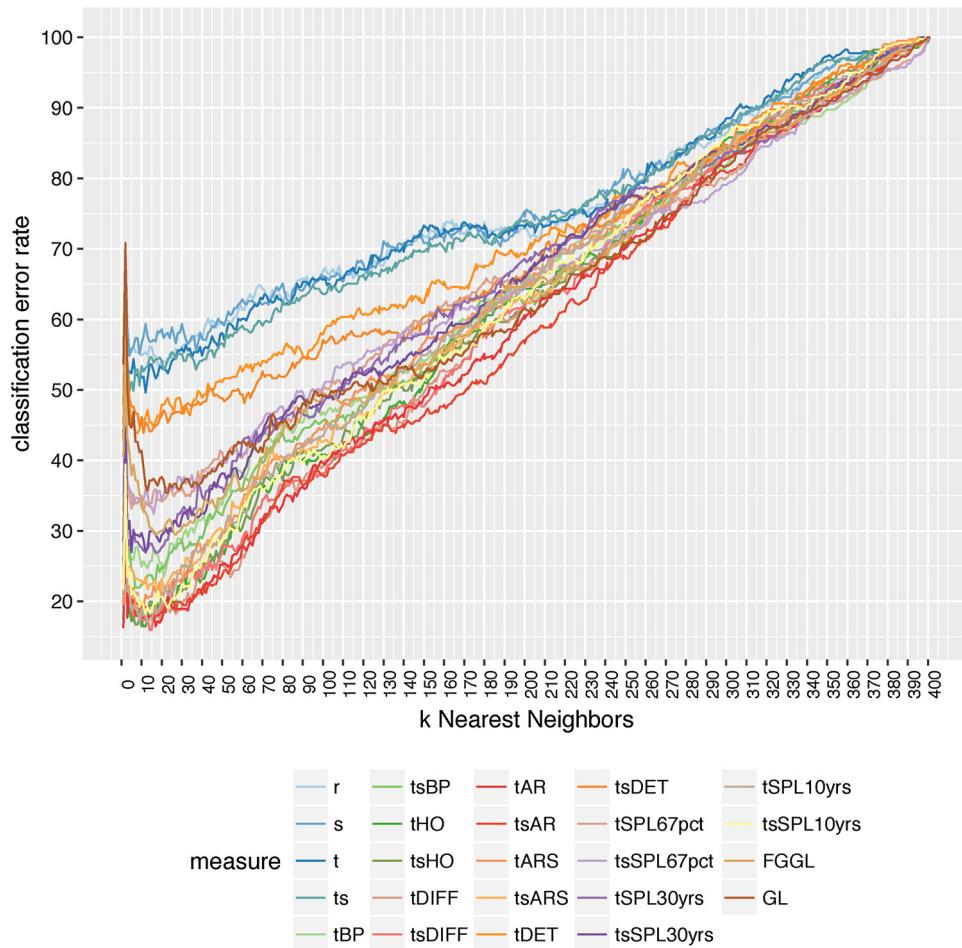


Fig. 8. Trend in classification error rate: Almost linear decrease until $k \approx 40$; afterwards, only slow further decrease or even stagnation (except for the jump at $k = 2$).

(*chw*). Also, a somewhat low rating score resulted for the site pair Sool-Hohwald (*how*) and Sool-Nübbannli (*nb*). These problematic cases formed a subset consisting of the sites at 1000–1106 m.

Generally, CER calculated on the complete data set decreased approximately linearly until $k \approx 40$. Afterwards, it decreased very slowly for most measures, or even stagnated (Fig. 8). An abrupt rise in CER occurred at stage $k = 2$ (Figs. 8 and A2) for all proximity measures. But overall, the stages $k = 40$ to $k = 1$ exhibited a relatively stable CER, and thus a relatively stable CSR, for most measures (Fig. 8). However, exactly in these stages, the sites *chw*, *kar* and *how* had a distinctly lower mean CSR than the others (Table 2), which coincided with the findings from the heat plot (Fig. 7). In addition, the site Gschwändwald (*gw*) attained a very low mean CSR although the heat plot indicated classification difficulties for the site pair *gw-kar* only. These four sites belong to the medium-elevation subsample, mostly misclassified to either sites on the same elevation, or sites that are located on the next lower elevational band (i.e., *fri*, *nb* and *sb*; Fig. 9). Moreover, some classification errors were caused by direct geographical neighbors (e.g., *how* series that were frequently misclassified as *nb* series; Fig. 9).

k-NN calculations excluding the sites *chw*, *how* and *kar* exhibited a rise in rating scores by an average of 3.4% (range 2.5% to 4.3%; blocks 5 and 6 in Table 6). In addition, \bar{CSR}_{open} rose by an average of 10.8% (range 4.8% to 14.1%). However, the rating scores and \bar{CSR}_{open} exhibited only a marginal further increase when the small sample of *gw* ($n = 15$) was excluded as well.

Classification errors occurred especially often in cases where the difference between the median of proximity values calculated for within-site comparisons and the median of proximity values calculated

for between-site comparisons was low (e.g., for the site pairs *kar-sb* and *how-nb*, Table 7). In contrast, the classification worked well between pairs of sites for which the difference of median proximity values calculated for between-site and within-site comparisons was high (e.g., site pairs *kar-sw* and *ww-rw*, Table 7).

4. Discussion

The main objective of this paper is to present a consistent method for evaluating the three key assumptions of dendro-provenancing stated in the Introduction. The k-NN classification was designed as a method whose success is highly dependent on the validity of the three key assumptions. Thus, the k-NN classification results need to be discussed before any inference on the validity of the dendro-provenancing assumptions can be drawn.

4.1. Suitability of different proximity measures for dendro-provenancing

Most proximity measures tested here involved preprocessing of the raw ring-widths prior to calculating a proximity value. Thus, the k-NN classification performance indicates which preprocessing method is most successful in enhancing the growth signal that is relevant for dendro-provenancing.

Correlation-based proximity measures that relied on high-pass filters performed best for all regional and elevation-specific subsets (i.e., *tHO*, *tsHO*, *tDIFF*, *tsDIFF*, *tAR*, *tsAR*, *tsSPL10yrs* and *tsSPL10yrs*). Interestingly, the differencing-based methods, i.e. *HO* and *DIFF*, were as efficient in enhancing the high-frequency variation as the very flexible

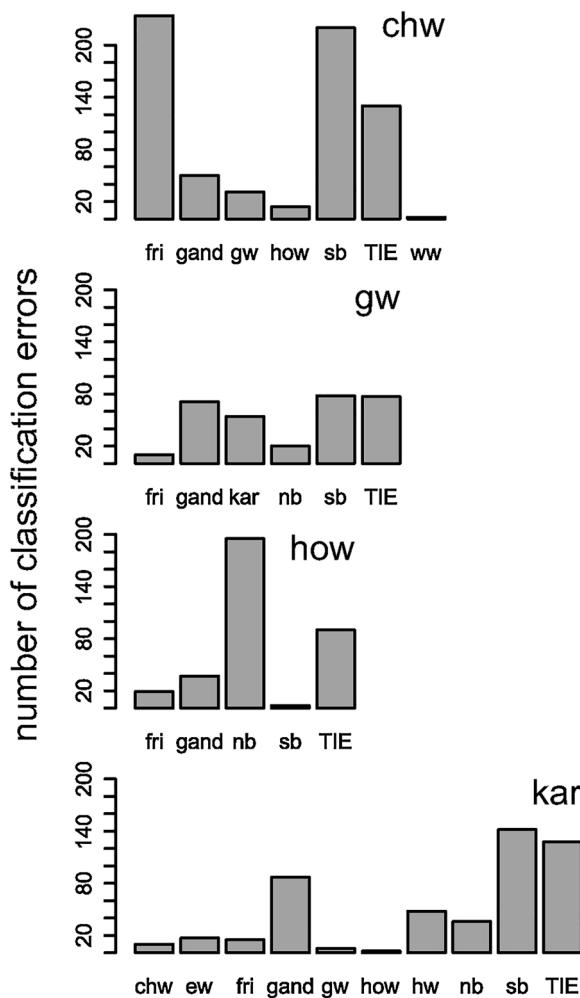


Fig. 9. Distribution of misclassified series per site. Upper right corner: site label of the site investigated (full names of sites see Table 1). The sites to which series have been misclassified are on the x-axis. The ‘TIE’ label denotes classification errors that resulted from a tie, i.e., no site label had a majority. The number of classification errors is on the y-axis and represents the sum of errors over the stages $k = 40$ till $k = 1$.

smoothing-splines of SPL10yrs or the auto-regressive modeling of AR. Using either the Pearson correlation coefficient or Spearman's rank correlation coefficient scarcely affected the classification performance. Although the Pearson correlation coefficient is known for its sensitivity to outliers (Edgell and Noon, 1984), this did not flaw the k-NN performance. In fact, the synchronous high frequency outliers (i.e., pointer years) to which the Pearson correlation coefficient is even more sensitive than the Spearman's rank correlation seem to occur often at the site level and thus are potentially provenancing-relevant.

The performance of ARS did not reach that of AR for the bulk of subsets, and the total data set. Hence, double detrending largely had a negative effect on classification performance. In contrast, although

direct autoregressive modeling of raw ring-width series probably resulted in biased autocorrelation coefficients (Monserud, 1986; Monserud and Yamaguchi, 1989; Cook and Kairiukstis, 1990; Stock and Watson, 2015), the autoregressive residual series as used by tAR and tsAR were successful here (cf. Supplementary Online Material). Nevertheless, future studies could profit from the individual single detrending of each series and the subsequent use of autoregressive moving-average (ARMA) models, which often are more parsimonious than AR models (Monserud, 1986; Monserud and Yamaguchi, 1989; Speer, 2010).

Measures implementing the BP transformation frequently reached quite high ratings, but not in all settings. Especially on the medium-elevation subsample, they performed worse than HO, DIFF, AR and SPL10yrs. Facing these alternatives, the BP transformation is considered a suboptimal choice for enhancing the provenancing relevant signal. Moreover, running mean transformations have been shown to cause spurious correlation (see Monserud, 1986). This most likely caused the lower k-NN classification performance of BP.

The relatively low performance of the sign-based measures FGGL and GL further supports that high-frequency filters in combination with correlation coefficients should be used for enhancing the signal relevant for dendro-provenancing. The rating for FGGL was close to that of tBP for many subsets. Moreover, FGGL, which takes the amplitude of growth changes into account (cf. Supplementary Online Material), performed much better than GL, which relies purely on sign-agreement. This again underlines the importance of synchronous high-frequency outliers for dendro-provenancing.

Even GL, with its rather low rating scores, performed better than the correlation-based methods that preserved low-frequency variation or completely lacked preprocessing (i.e., DET, ts, t, r, s). Clearly, preserving the low-frequency variation did not pay off in terms of classification success, as demonstrated by the worsening performance of the sequence SPL10yrs, SPL30yrs, SPL67pct, DET, ts, t, r, and s. However, the 1180–1196 m subsample featured completely diverging persistence patterns, but it consisted of just two sites (gw and gand). In this exceptional case, the high-pass filtering removed the diverging persistence patterns and thus removed the feature that differed most between the ring-width series of both sites. Nevertheless, high-pass filter based methods performed just a little worse on this subsample than low-frequency tolerant methods. Moreover, persistence patterns, like those differentiating the sites of the 1180–1196 m subsample, lost their relevance for classification as soon as the data set was extended. Thus, in the data set studied, similar medium- to low-frequency patterns were not as site-specific as similarities in high-frequency variation. Correlation coefficients can only detect common trends and medium-frequency growth patterns in the time domain when the preprocessing preserves these patterns. At the same time, exactly these patterns (i.e., trends and medium frequency patterns) obscure the high-frequency correlation. Hence, other proximity measures are probably more suitable to reveal persistent growth variations that could be relevant for dendro-provenancing. This may be accomplished for example by measures that account for distances in the frequency domain or calculate distances between auto-regressive model coefficients (Hennig, 2016).

Table 7

Median t -values (tHO) for within-site and between-site pairwise comparisons. Due to space constraints, only these exemplary results are presented and are limited to tHO, as the findings were quite similar for all high-performance proximity measures (i.e., tHO, tsHO, tAR, tsAR, tDIFF, tsDIFF, tsSPL10yrs and tsSPL10yrs).

site A	site B	within A	within B	between AB	within A - within B	within A - between AB	within B - between AB
kar	sb	2.97	4.31	2.57	-1.34	0.40	1.74
how	nb	5.91	7.82	5.33	-1.92	0.57	2.49
kar	sw	2.97	4.05	1.71	-1.08	1.27	2.34
ww	rw	6.87	6.86	5.79	0.01	1.08	1.07

4.2. Regional and elevation-specific differences in classification performance

The regional subsets consist of sites spread across all elevational belts of a region, whereas climate-growth relationships are known to be most pronounced along elevational gradients (Frank and Esper, 2005; Savva et al., 2006; Eissing and Dittmar, 2011; Dittmar et al., 2012; King et al., 2013; Kolář, Čermák et al., 2017; Lyu et al., 2017). Hence, differences in between-site growth variation within a regional subset are likely attributable to differences in elevation. As the classification performance was similar for all regional subsets, the effect of elevation on the between-site growth variation seems to be equal across the regions.

Within the elevational subsets, which comprise sites of all regions, series were successfully classified to their original sites by correlation-based measures that involve high-pass filtering. Thus, the high-frequency variation between sites of the same elevation differs strongly enough in most cases, except for the medium-elevation subsample.

4.3. Analysis of classification errors

Classification difficulties arose for all measures when the classification was based on two NN. In such instances just one off-site-NN needed to be among the pair of classifier-NN to cause a tie. Since the classification error rate increased dramatically at $k = 2$, many of the ring-width series must have had at least one off-site NN among their two NN.

Besides this $k = 2$ classification problem, notable difficulties were caused by medium-elevation sites that were misclassified to sites at 829–1198 m. Investigating the within-site and between-site median THO values showed that the lack of signal difference between on-site and off-site series increased the frequency of classification errors. Thus, between-site variability was seemingly less pronounced between upper low-elevation and medium-elevation than at other elevational belts. Within the upper low-elevation and medium-elevation belt, classification errors were not exclusively confined to sites that exhibited similar site-factor combinations and/or were located in direct geographical proximity (such as the site pairs *kar-gw*, *how-nb*, *chw-sb* and *chw-fri*; cf. Table 1 and Figs. 7, 9). Problems also arose between distant sites with differing site-factor combinations (such as *kar-sb*, *kar-gand*; cf. Table 1 and Figs. 7, 9). This contrasts the findings of Bridge (2000) who found good matches between series from distant sites provided that they share key site factors.

Between-site high-frequency variability is thought to be controlled primarily by climatic factors and weather conditions (Fritts, 1976). Studies of the climate-growth relationship of spruce in the Bavarian forest (Dittmar and Elling, 1999; Wilson and Hopfmüller, 2001) may offer some preliminary explanation for the classification problems encountered here, as discussed below.

Wilson and Hopfmüller (2001) found that growth below 680 m was predominantly controlled by moisture availability. In their study, no significant correlations between growth and climate parameters were found for sites between 780 and 970 m. Even at ≥ 1070 m, site chronologies still exhibited a relatively weak climate signal. Unexpectedly, temperature had no dominant control upon growth at these higher sites.

Dittmar and Elling (1999) also investigated climate-growth relationships, but in their study the sites where no dominant controls on growth could be identified lay somewhat lower than those reported by Wilson and Hopfmüller (2001), i.e. at elevations between 600 and 800 m, and above 800 m temperature gradually became the dominant factor limiting growth. Below 600 m, water supply was found to predominantly control growth, which agreed with the finding by Wilson and Hopfmüller (2001).

Thus, although their findings agreed in some cases, the two studies found somewhat different elevational belts to be most challenging for

interpreting climate-growth relationships. Hence, more research is needed to determine the elevational belt where such complex interactions occur within the region studied here. However, Wilson and Hopfmüller (2001) found that sites between 780–970 m shared much of the year-to-year variability with chronologies located at ≥ 1070 m (of these chronologies, four are located at 1070–1230 m and two are located at 1325–1420 m). This elevational belt with lower between-site variability and no significant correlations between growth and meteorological data matches well with the elevational belt where k-NN classification problems occurred here (829–1198 m).

5. Conclusions

Scissor-plot rating of k-NN offers a consistent method for evaluating the three fundamental dendro-provenancing assumptions prior to determining the provenance of wood. The successful k-NN classification of ring-width series to their original site indicates that growth of spruce diverges sufficiently within the study area to allow for the rise of site-specific high-frequency growth patterns. Thus, the first assumption of dendro-provenancing seems to be adequate for the data set studied here. Moreover, the results imply that similar growth can be quantified by statistical measures of proximity (the second assumption), and that best proximity values are closely correlated with geographical neighbors (except for classifications using $k = 2$), which is the third key assumption of dendro-provenancing.

The between-site variation of the high-frequency signal is paramount for a successful classification. Consequently, correlation-based measures that rely on high-pass filters performed best (tHO, tsHO, tDIFF, tsDIFF, tAR, tsAR, tsSPL10yrs and tsSPL10yrs). Preserving the mid- to low-frequency variation did not have any positive effect on classification performance.

Classification problems arose where the difference between within-site and between-site signal was small. In the data set presented here, such small differences were predominantly limited to medium-elevation sites between 1000 and 1198 m. Moreover, the low performance of the classification using $k = 2$ illustrated that relying on a single setting for k may easily lead to unreliable classifications. The climatic and possibly ecological causes underlying the between-site growth variation need to be analyzed further.

Using scissor plots, the stability of classifications can be investigated. In the future, scissor plots could also be evaluated for k-NN classifications based on tree-ring variables other than ring-width data, such as blue intensity, density, or stable isotopes.

Acknowledgements

I thank N. Bleicher for comments, discussions and suggestions; H. Bugmann for comments, discussions and suggestions and providing me with meteorological data; Ph. Della Casa and H. Bugmann for supervising my PhD thesis; and T. Eissing and two anonymous reviewers for additional comments and discussions; also, I thank M. Bolliger, K. Emmenegger, S. Forabosco, F. Gut, L. Gut, T. Jagouls, N. Lengacher, J. Schmidt, T. Wäckerle, F. Walder and S. Wicki for support in the field work. This study was financially supported by the Swiss National Science Foundation, grant no. P0ZHP1_162299.

Appendix A. Supplementary Data

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.dendro.2018.09.008>.

References

- Arlot, S., Celisse, A., 2010. A survey of cross-validation procedures for model selection. *Stat. Surv.* 4, 40–79.
- Baillie, M.G.L., Pilcher, J.R., 1973. A simple cross-dating programme for tree-ring research.

- Tree-Ring Bull. 33, 7–14.
- Best, D.J., Roberts, D.E., 1975. Algorithm AS 89: the upper tail probabilities of Spearman's rho. *Appl. Stat.* 24, 377–379.
- Boschetti-Maradi, A., Kontic, R., 2012. Möglichkeiten und Schwierigkeiten dendrochronologischer Untersuchungen in Mittelalterarchäologie und Bauforschung. *Mitteilungen der Deutschen Gesellschaft für Archäologie des Mittelalters und der Neuzeit* 24, 49–60.
- Bridge, M., 2000. Can dendrochronology be used to indicate the source of oak within Britain? *Vernac. Archit.* 31, 67–72.
- Bridge, M., 2012. Locating the origins of wood resources: a review of dendroprovenancing. *J. Archaeol. Sci.* 39, 2828–2834.
- Brockwell, P.J., Davis, R.A., 1996. *Introduction to Time Series and Forecasting*. Springer Texts in Statistics. Springer, Berlin and Heidelberg.
- Bronaugh, D., 2018. climdex.pcic: PCIC Implementation of Climdex Routines. R package version 1.1-9.
- Buras, A., Wilmking, M., 2015. Correcting the calculation of Gleichläufigkeit. *Dendrochronologia* 34, 29–30.
- Cook, E.R., 1981. The smoothing spline: a new approach to standardizing forest interior tree-ring width series for dendroclimatic studies. *Tree-Ring Bull.* 41, 45–53.
- Cook, E.R., 1985. A time series approach to tree ring standardization. University of Arizona, Tucson (Dissertation).
- Cook, E.R., Kairikstis, L.A. (Eds.), 1990. *Methods of Dendrochronology*. Kluwer, Dordrecht.
- Cover, T., Hart, P., 1967. Nearest neighbor pattern classification. *IEEE Trans. Inf. Theory* 13, 21–27.
- Daly, A., Nymoen, P., 2008. The Bøle ship, Skien, Norway—Research history, dendrochronology and provenance. *Int. J. Nautic. Archaeol.* 37, 153–170.
- Dittmar, C., Eissing, T., Rothe, A., 2012. Elevation-specific tree-ring chronologies of Norway spruce and Silver fir in Southern Germany. *Dendrochronologia* 30, 73–83.
- Dittmar, C., Elling, W., 1999. Jahrringbreite von Fichte und Buche in Abhängigkeit von Witterung und Höhenlage. *Forstwissenschaftliches Centralblatt* 118, 251–270.
- Drake, B.L., 2018. Source & sourceability: towards a probabilistic framework for dendroprovenance based on hypothesis testing and Bayesian inference. *Dendrochronologia* 47, 38–47.
- Eckstein, D., Wrobel, S., 2007. Dendrochronological proof of origin of historic timber – retrospect and perspectives. In: Haneca, K., Verheyden, A., Beeckman, H., Gärtnner, H., Helle, G., Schleser, G. (Eds.), *TRACE – Tree Rings in Archaeology, Climatology and Ecology*, Vol. 5: Proceedings of the Dendrosymposium 2006, April 20th–22nd 2006, Tervuren, Belgium, pp. 8–20.
- Edgell, S.E., Noon, S.M., 1984. Effect of violation of normality on the t test of the correlation coefficient. *Psychol. Bull.* 95, 576–583.
- Eissing, T., 2007. Zum Problem der Interpretation dendrochronologischer Datierung bei Flossholz am Beispiel Bamberg's. In: *für Hausforschung*, A. (Ed.), Spuren der Nutzung in historischen Bauten. Jonas, Marburg. *Jahrbuch für Hausforschung*, 54, pp. 23–36.
- Eissing, T., Dittmar, C., 2011. Timber transport and dendroprovenancing in Thuringia and Bavaria. In: Fraiture, P. (Ed.), *Treerings, Art and Archaeology*, pp. 137–149.
- Everitt, B.S., Sabine, L., Morven, L., Stahl, D., 2011. *Cluster Analysis*. Wiley Series in Probability and Statistics 848, 5th edition. Wiley, Chichester.
- Fowler, A., Bridge, M., 2015. Mining the British Isles oak tree-ring data set. Part A: Rationale, data, software, and proof of concept. *Dendrochronologia* 35, 24–33.
- Frank, D., Esper, J., 2005. Characterization and climate response patterns of a high-elevation, multi-species tree-ring network in the European Alps. *Dendrochronologia* 22, 107–121.
- Fritts, H.C., 1976. *Treerings and Climate*. Academic Press, London, New York, San Francisco.
- Garcia-Gonzalez, I., 2008. Comparison of different distance measures for cluster analysis of tree-ring series. *Tree-Ring Res.* 64, 27–37.
- Grissino-Mayer, H.D., 2001. Evaluating crossdating accuracy: a manual and tutorial for the computer program COFECHA. *Tree-Ring Res.* 57, 205–221.
- Haneca, K., Wasny, T., van Acker, J., Beeckman, H., 2005. Provenancing Baltic timber from art historical objects: success and limitations. *J. Archaeol. Sci.* 32, 261–271.
- Hellmann, L., Tegel, W., Geyer, J., Kirdyanov, A.V., Nikolaev, A.N., Eggertsson, O., Altman, J., Reinig, F., Morganti, S., Wacker, L., Büntgen, U., 2017. Dendro-provenancing of Arctic driftwood. *Q. Sci. Rev.* 162, 1–11.
- Hennig, C.M., Meilă, M., Murtagh, F., Rocci, R., 2016. *Handbook of cluster analysis*. Chapman & Hall/CRC handbooks of modern statistical methods. CRC Press, Taylor & Francis Group, Boca Raton.
- Hollander, M., Wolfe, D.A., 1973. *Nonparametric Statistical Methods*. Wiley Series in Probability and Statistics. John Wiley and Sons Ltd., New York.
- Hollstein, E., 1980. Mitteleuropäische Eichenchronologie. Trierer dendrochronologische Forschungen zur Archäologie und Kunstgeschichte 11, Philipp von Zabern, Mainz am Rhein.
- Huber, B., 1943. Über die Sicherheit jahrringchronologischer Datierung. Holz als Roh- und Werkstoff 6 263 ff.
- Jansma, E., Haneca, K., Kosian, M., 2014. A dendrochronological reassessment of three Roman vessels from Utrecht (the Netherlands): Evidence of inland navigation between the lower-Scheldt region in Gallia Belgica and the limes of Germania inferior. *J. Archaeol. Sci.* 50, 1–13.
- von Jazelewitsch, W., 1948. Über die Möglichkeiten einer jahrringchronologischen Individualdiagnose von Bäumen mit Beiträgen zur Methodik der Jahrringforschung. Ludwig-Maximilians-Universität, München (Dissertation).
- Kendall, M.G., Gibbons, J.D., 1990. *Rank Correlation Methods*, 5th ed. Edward Arnold, London.
- King, G.M., Gugerli, F., Fonti, P., Frank, D.C., 2013. Tree growth response along an elevational gradient: climate or genetics? *Oecologia* 173, 1587–1600.
- Kolář, T., Čermák, P., Trnka, M., Žid, T., Rybníček, M., 2017. Temporal changes in the climate sensitivity of Norway spruce and European beech along an elevation gradient in Central Europe. *Agric. For. Meteorol.* 239, 24–33.
- Lyu, L., Suvanto, S., Nöjd, P., Henttonen, H.M., Mäkinen, H., Zhang, Q.B., 2017. Tree growth and its climate signal along latitudinal and altitudinal gradients: comparison of tree rings between Finland and the Tibetan Plateau. *Biogeosciences* 14, 3083–3095.
- Monserud, R.A., 1986. Time-series analysis of tree-ring chronologies. *For. Sci.* 32, 349–372.
- Monserud, R.A., Yamaguchi, D.K., 1989. Comments on “Cross-dating methods in dendrochronology” by Wigley et al. *J. Archaeol. Sci.* 16, 221–224.
- Neuwirth, B., Esper, J., Schweingruber, F.H., Winiger, M., 2004. Site ecological differences to the climatic forcing of spruce pointer years from the Lötschental, Switzerland. *Dendrochronologia* 21, 69–78.
- Neuwirth, B., Schweingruber, F.H., Winiger, M., 2007. Spatial patterns of central European pointer years from 1901 to 1971. *Dendrochronologia* 24, 79–89.
- Pearson, K., 1895. Note on regression and inheritance in the case of two parents. *Proc. R. Soc. Lond. U.S.A.* 58, 240–242.
- R Core Team, 2017. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rolland, C., Desplanque, C., Michalet, R., Schweingruber, F.H., 2000. Extreme tree rings in spruce (*Picea abies* [L.] Karst.) and fir (*Abies alba* Mill.) stands in relation to climate, site, and space in the Southern French and Italian Alps. *Arctic Antarctic Alpine Res.* 32, 1.
- Savva, Y., Oleksyn, J., Reich, P.B., Tjoelker, M.G., Vaganov, E.A., Modrzynski, J., 2006. Interannual growth response of Norway spruce to climate along an altitudinal gradient in the Tatra Mountains, Poland. *Trees* 20, 735–746.
- Schmitt, I., 2006. Ähnlichkeitssuche in Multimedia-Datenbanken: Retrieval, Suchalgorithmen und Anfragebehandlung. Oldenbourg, München.
- Speer, J.H., 2010. *Fundamentals of Tree-ring Research*. University of Arizona Press, Tucson.
- Stock, J.H., Watson, M.W., 2015. *Introduction to Econometrics*. Always Learning, 3rd global edition. Pearson, Boston.
- Thornton, P.E., Running, S.W., White, M.A., 1997. Generating surfaces of daily meteorological variables over large regions of complex terrain. *J. Hydrol.* 190, 214–251.
- Wazny, T., 2002. Baltic timber in Western Europe – an exciting dendrochronological question. *Dendrochronologia* 20, 313–320.
- Wigley, T., Jones, P.D., Briffa, K.R., 1987. Cross-dating methods in dendrochronology. *J. Archaeol. Sci.* 14, 51–64.
- Wilson, R., Hopfmüller, M., 2001. Dendrochronological investigations of Norway spruce along an elevational transect in the Bavarian Forest, Germany. *Dendrochronologia* 19, 67–79.

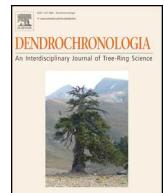
(2) Artikel II

No systematic effects of sampling direction on climate-growth relationships in a large-scale, multi-species tree-ring data set.

Urs Gut, Mátyás Árvai, Szymon Bijak, J. Julio Camarero, Anna Cedro, Roberto Cruz-García, Balázs Garamszegi, Andrew Hacket-Pain, Andrea Hevia, Weiwei Huang, Miriam Isaac-Renton, Ryszard J. Kaczka, Marko Kazimirović, Wojciech Kędziora, Zoltán Kern, Marcin Klisz, Tomáš Kolář, Michael Körner, Veronica Kuznetsova, David Montwé, Any Mary Petritan, Ion Catalin Petritan, Lenka Plavcová, Romy Rehschuh, Eva Rocha, Michal Rybníček, Raúl Sánchez-Salguero, Jens Schröder, Niels Schwab, Branko Stajić, Robert Tomusiak, Martin Wilmking, Ute Sass-Klaassen, Allan Buras

Dieser Artikel wurde veröffentlicht in:

Gut, Urs, et al., (2019). No systematic effects of sampling direction on climate-growth relationships in a large-scale, multi-species tree-ring data set. *Dendrochronologia*, 57, 1–9. 10.1016/j.dendro.2019.125624.



No systematic effects of sampling direction on climate-growth relationships in a large-scale, multi-species tree-ring data set



Urs Gut^{a,b,*}, Mátyás Árvai^c, Szymon Bijak^d, J. Julio Camarero^e, Anna Cedro^f, Roberto Cruz-García^g, Balázs Garamszegi^h, Andrew Hacket-Painⁱ, Andrea Hevia^j, Weiwei Huang^{k,l}, Miriam Isaac-Renton^m, Ryszard J. Kaczkaⁿ, Marko Kazimirović^o, Wojciech Kędziora^p, Zoltán Kern^q, Marcin Klisz^r, Tomáš Kolář^{s,t}, Michael Körner^u, Veronica Kuznetsova^v, David Montwé^w, Any Mary Petritan^x, Ion Catalin Petritan^y, Lenka Plavcová^z, Romy Rehschuh^A, Eva Rocha^B, Michal Rybníček^{s,t}, Raúl Sánchez-Salguero^C, Jens Schröder^u, Niels Schwab^D, Branko Stajić^o, Robert Tomusia^d, Martin Wilmking^g, Ute Sass-Klaassen^E, Allan Buras^{E,F}

^a Institut für Archäologie, Fachbereich Prähistorische Archäologie, University of Zurich, 8006, Zurich, Switzerland

^b Department of Environmental System Science, Institute of Terrestrial Ecosystems, Forest Ecology, ETH Zurich, 8092, Zurich, Switzerland

^c Institute for Soil Sciences and Agricultural Chemistry, MTA Centre for Agricultural Research, 1022, Budapest, Hungary

^d Dendrometry and Forest Productivity, Warsaw University of Life Sciences – SGGW, 02-776, Warszawa, Poland

^e Pyrenean Institute of Ecology, Spanish Research Council, CSIC, 50059, Zaragoza, Spain

^f Department of Geosciences, Szczecin University, 70–383, Szczecin, Poland

^g Landscape Ecology and Ecosystem Dynamics, Greifswald University, 17487, Greifswald, Germany

^h Forest Research Institute, National Agricultural Research and Innovation Centre, 9600, Sárvar, Hungary

ⁱ Department of Geography and Planning, University of Liverpool, L69 7ZT, Liverpool, United Kingdom

^j Dept. de Ciencias Agroforestales, Escuela Técnica Superior de Ingeniería, Universidad de Huelva, 21819, La Rábida, Spain

^k Bamboo Research Institute, Nanjing Forestry University, 210037, Nanjing, China

^l Department of Geosciences and Natural Resource Management, University of Copenhagen, 1958, Frederiksberg C, Denmark

^m Canadian Wood Fibre Centre, Natural Resources Canada, V8Z 1M5, Victoria, Canada

ⁿ Faculty of Earth Sciences, University of Silesia, 41-200, Sosnowiec, Poland

^o Faculty of Forestry, University of Belgrade, 11000, Belgrade, Serbia

^p Forest Management Planning, Warsaw University of Life Sciences – SGGW, 02-776, Warszawa, Poland

^q Institute for Geological and Geochemical Research, Research Centre for Astronomy and Earth Sciences, MTA, 1112, Budapest, Hungary

^r Department of Silviculture and Forest Tree Genetics, Forest Research Institute, 05-090, Raszyn, Poland

^s Department of Wood Science, Mendel University, 613 00, Brno, Czech Republic

^t Global Change Research Institute of the Czech Academy of Sciences, 603 00, Brno, Czech Republic

^u Faculty of Forest and Environment, Eberswalde University for Sustainable Development, 16225, Eberswalde, Germany

^v Glaciology, Institute of Geography, Russian Academy of Sciences, 119017, Moscow, Russia

^w Centre for Forest Biology, University of Victoria, V8W 2Y2, Victoria, Canada

^x Forest Ecology, National Institute for Research and Development in Forestry “Marin Drăcea”, 500040, Brasov, Romania

^y Sylviculture and Forest Engineering, University of Transylvania, 500123, Brasov, Romania

^z Department of Biology, University of Hradec Králové, 500 03, Hradec Králové, Czech Republic

^A Institute of Meteorology and Climate Research – Atmospheric Environmental Research (IMK-IFU), Karlsruhe Institute of Technology (KIT), 82467, Garmisch-Partenkirchen, Germany

^B Department of Physical Geography, Stockholm University, 106 91, Stockholm, Sweden

^C Dept. de Sistemas Físicos, Químicos y Naturales, Universidad Pablo de Olavide, 41013, Sevilla, Spain

^D Physical Geography, Center for Earth System Research and Sustainability (CEN), Universität Hamburg, 20146, Hamburg, Germany

^E Forest Ecology and Forest Management, Wageningen University and Research, 6700 PB, Wageningen, Netherlands

^F Land Surface-Atmosphere Interactions, Technische Universität München, 85354, Freising, Germany

* Corresponding author at: Universität Zürich, Institut für Archäologie, Fachbereich Prähistorische Archäologie, Karl Schmid-Strasse 4, 8006 Zürich, Switzerland.
E-mail address: u.gut@posteo.de (U. Gut).

ARTICLE INFO

Keywords:
 Tree-rings
 Directional growth
 Climate signal
 Dendro-provenancing
 Principal Component Gradient Analysis
 Correlation analysis

ABSTRACT

Ring-width series are important for diverse fields of research such as the study of past climate, forest ecology, forest genetics, and the determination of origin (dendro-provenancing) or dating of archaeological objects. Recent research suggests diverging climate-growth relationships in tree-rings due to the cardinal direction of extracting the tree cores (i.e. direction-specific effect). This presents an understudied source of bias that potentially affects many data sets in tree-ring research.

In this study, we investigated possible direction-specific growth variability based on an international (10 countries), multi-species (8 species) tree-ring width network encompassing 22 sites. To estimate the effect of direction-specific growth variability on climate-growth relationships, we applied a combination of three methods: An analysis of signal strength differences, a Principal Component Gradient Analysis and a test on the direction-specific differences in correlations between indexed ring-widths series and climate variables.

We found no evidence for systematic direction-specific effects on tree radial growth variability in high-pass filtered ring-width series. In addition, direction-specific growth showed only marginal effects on climate-growth correlations. These findings therefore indicate that there is no consistent bias caused by coring direction in data sets used for diverse dendrochronological applications on relatively mesic sites within forests in flat terrain, as were studied here. However, in extremely dry, warm or cold environments, or on steep slopes, and for different life-forms such as shrubs, further research is advisable.

1. Introduction

Tree-ring records provide valuable data for various scientific disciplines. For example, tree-rings are used in paleoclimatology, as a proxy for reconstructing past climate (Hughes et al., 2011); in ecology, for investigating stand dynamics (Amoroso et al., 2017; Schweingruber, 1996); and in history and archaeology, for dating of artifacts and wooden construction elements as well as for analyzing past usage of timber resources (Bleicher and Harb, 2015; Eissing and Dittmar, 2011).

Tree growth is dependent on environmental, climatic and biotic factors (Cook and Kairiukstis, 1990). Through careful sampling design and selection of appropriate statistical methods, certain growth signals in ring-width series may be selectively enhanced (Fritts, 1976; Sullivan and Csank, 2016). In dendro-climatology, for example, the climatically-related environmental signal is of primary importance. Thus, sampling is commonly conducted on trees growing at the species distributional margins where the desired climatic factor is assumed to limit tree-growth, for example, growing-season temperature in cold environments or precipitation in dry environments (Fritts, 1976; Klesse et al., 2018). In addition, statistical detrending is used to remove the age-size-related trend (Cook and Kairiukstis, 1990; Melvin and Briffa, 2008; Peters et al., 2015). To amplify growth signals shared by a tree population and averaging out individual-specific noise, mean site tree-ring width chronologies are calculated (Cook and Kairiukstis, 1990). Often, this approach leads to absolutely higher correlations between climate variables and master chronologies in comparison to climate correlations at the individual tree level (Carrer, 2011; Galván et al., 2014).

A recent study by Fang et al. (2015) suggested the potential for inadvertent sampling bias due to commonly-used sampling procedures such as coring individual trees from random or 'ad hoc' radial directions. Their results provided evidence for diverging climate-growth relationships between ring-width series from the same tree cored from different cardinal directions for *Pinus tabuliformis* Carrière and *Picea purpurea* Mast. Thus, mean site chronologies might enhance an already biased climate signal if they were based on a low replication and/or samples that were taken with a systematic preference for an azimuth coring direction. For example, most standard sampling protocols recommend extracting tree cores at positions where the trunk is perpendicular to the slope and parallel to the contour line (Speer, 2010; Yang et al., 2018). Such sampling protocols seek to avoid reaction wood. At the same time, certain cardinal directions of the stem are favored systematically. Consequently, if direction-specific climate-growth relationships exist, conventional sampling protocols seeking to avoid reaction wood may unintentionally produce biased chronologies. Therefore, potentially all dendro-ecological and dendro-climatological data sets that were obtained following such standard sampling protocols might be affected.

Increment cores taken from historical and archaeological contexts may be particularly sensitive to potential direction-specific bias. This is because, for historical and archaeological wood, the original orientation of the corresponding tree can generally not be determined. Moreover, multiple cores from different – preferably orthogonal – directions are rarely extractable due to the limited accessibility of construction timber in buildings and the artefact-status of wooden archaeological objects. Consequently, if direction-specific effects on growth variability and climate-growth relationships exist, biased chronologies could result. Hence, investigations of direction-specific effects are needed to quantify the bias possibly affecting historic and archaeological data sets that form the backbone of millennia-long chronologies used for climate reconstructions and dendro-provenancing (Bridge, 2012; PAGES 2k Consortium, 2013).

The within- and between-tree sources of noise have been studied from the onset of dendrochronological research (Fritts, 1976). However, systematic investigations of intra-individual growth variability in trees are rare, partly because of the considerable effort needed for collecting sufficient samples to assess tree-growth along the whole stem (Babst et al., 2018; Duncker and Spiecker, 2008). In contrast, growing evidence of direction-specific radial growth was recently documented for several shrub species (Buras and Wilmking, 2014; Shetti et al., 2018; Yang et al., 2018). Moreover, Gričar et al. (2006) showed that heating and cooling of stem sections can affect cambial activity and cell differentiation in *Picea abies* (L.) H. Karst, which may result in direction-specific growth variability if temperature differences between different stem-parts would prevail over several years (but see Buras and Wilmking, 2014 for an elaborated discussion). Changing climatic responses have been documented in stem disk mean chronologies taken at different heights (Chhin et al., 2010). Besides the investigation of Fang et al. (2015), to our knowledge, there exists no other study that reported direction-specific climate-growth relationships in trees among samples taken at the commonly-sampled height of 1.3 m (diameter at breast height).

Potential direction-specific climate-growth relationships could prove problematic, especially if the year-to-year variability was affected. In addition to being the focus of climatological research, high-frequency growth signals – representing the year-to-year variability – are critical for dendro-provenancing analyses (Gut, 2018). Diverging direction-specific climate-growth relationships, as observed by Fang et al. (2015), thus may result in direction-specific biases that could affect I) the chronologies used for millennia-long climatic reconstructions (PAGES 2k Consortium, 2013), II) the local reference chronologies used for dendro-provenancing and III) the site chronologies established for dendro-ecological studies.

Given the possibly wide-ranging impact on tree-ring research, we here examine the potential for differences in growth variability and

climate signals due to the azimuth orientation of tree increment coring. We use tree increment cores from an international tree-ring network that was designed to answer the following research questions:

1. Is there evidence for direction-specific high-frequency growth variability in our data set?
2. Are systematic direction-specific differences in climate-growth relationships present in our data set?

2. Material and methods

2.1. Tree-ring and climate data

To assess possible direction-specific climate signals across different sites and species, the last author of this study (A. Buras) released an international call for contributing tree-ring data in autumn 2017 via the Association of Tree-Ring Research (ATR). The sampling followed a uniform design to reduce investigator bias and ensure that cores were sampled systematically regarding the cardinal direction of the stem. The specifications of the sampling design were: i) 15 trees per site, ii) sampled for two cores from the southern and eastern direction of the stem, iii) collected within a forest, not closer than 100 m to the nearest forest edge of iv) homogeneous, ideally mono-specific forests on v) preferably flat terrain to avoid any potential confounding effects from reaction wood in trees growing on slopes. In addition, for each sampled tree, tree height and diameter at breast height (DBH) were recorded.

All sites met the criteria ii, iii, and v, but at seven sites (THPA, GBPS, REPS, RAPS, AVPN, DPPS, UKPA) less than 15 (minimum twelve) trees were sampled for east- and south-facing increment cores due to the occurrence of wounds or other irregularities and at five sites (OMPS, GBPS, NIPS, CRPG, TLPM) few individuals (in total less than 10% of standing trees) of other tree species were observed. However, since these sites did not reveal differing results in our analyses, we assume possible associated effects to be negligible. All European sites represent managed forests, but the two sites from Canada were located in naturally-grown forests. No data on tree height exists for the sites HCQC and HEQP. For RAPS, DBH was not measured.

Sample preparation, cross-dating and measurement of total ring-widths, as well as quality control of ring-width chronologies, were done

following standard dendrochronological procedures (Cook and Kairiukstis, 1990; Grissino-Mayer, 2001; Speer, 2010). Following this protocol, the final data set encompassed 22 sites covering 8 tree species (Table 1, Figs. 1, 2).

For the series of monthly temperature averages, the 0.5°-gridded CRU TS3.10 data set was used (Harris et al., 2014). For monthly precipitation sums, the 0.5°-gridded GPCC data set was used (Schneider et al., 2016). Tree-ring, climate and tree meta data were analyzed with the statistical software R version 3.5.1 (R Core Team, 2018).

2.2. Preprocessing and signal strength statistics

Given our research question, high-frequency growth variability was the focus of this study. Hence, autoregressive models were fit to the individual ring-width series as derived from the increment cores. The optimal model was chosen as the model that minimized Akaike's Information Criterion (i.e. the default settings of the autoregressive detrending method implemented in the 'dplR' package, Bunn et al., 2018; Bunn, 2008). The resulting series of autoregressive residuals were divided by the mean of the residuals to obtain white noise residual ring-width indices (RWI) with a mean of 1. As a measure of direction-specific signal strength, we computed the mean inter-series Gleichläufigkeit (glk, Buras and Wilmking, 2015; Eckstein and Bauch, 1969) as well as the mean inter-series correlation (r , Wigley et al., 1984) pooled by southern and eastern cores. To avoid possible series-length effects, all computations were conducted over the common overlap period of the respective site chronologies. Both statistics were calculated over all possible pairwise comparisons of a respective cardinal direction subset. To bolster robustness, the Spearman's rank sum correlation coefficient was used for all correlation analyses in this study (Best and Roberts, 1975; Hollander et al., 2015).

Besides direction-specific between-series synchronicity, we considered two additional methodological approaches to assess possible direction-specific growth signals: i) direction-specific aggregation of growth signals, and ii) direction-specific climate-growth relationships, which are described in sections 2.3 and 2.4, respectively.

Table 1

Meta data summary for all sites and species included in the study. Sites are sorted by country. Latitude and longitude in decimal degrees. N: Number of trees for which an eastern and a southern core was sampled; 1 st yr, Last yr: First and last year of the common overlap period shared by all ring-width series of a site; M.s.length: Mean series length in years (calculated across eastern and southern cores); M.DBH: Mean diameter at breast height in cm; M.height: Mean tree height in m. Temp: Mean annual temperature in °C; Prec: Mean yearly precipitation sums in mm.

Country	Site	Abbrev.	Species	Latitude	Longitude (-W, +E)	N	1 st yr	Last yr	M.s.length	M.DBH	M.height	Temp	Prec
Canada	Cline River	CRPG	<i>Picea glauca</i> (Moench) Voss	52.000	-116.507	15	1940	2017	121	34	21	-2.1	664
Canada	Thetis Lake	TLPM	<i>Pseudotsuga menziesii</i> (Mirb.) Franco	48.462	-123.465	15	1928	2017	129	72	33	9.3	720
Czech Republic	Košetice	KSPA	<i>Picea abies</i> (L.) H.Karst.	49.566	15.091	15	1962	2017	75	45	29	7.8	626
Czech Republic	Soběšice	SOQP	<i>Quercus petraea</i> (Matt.) Liebl.	49.250	16.610	15	1984	2014	87	40	17	8.1	569
Germany	Garmisch	GAPA	<i>Picea abies</i> (L.) H.Karst.	47.494	11.065	15	1963	2017	97	51	36	3.1	1051
Germany	Niederhaverbek	NIPS	<i>Pinus sylvestris</i> L.	53.132	9.875	15	1984	2017	40	29	20	8.7	743
Germany	Tharandt	THPA	<i>Picea abies</i> (L.) H.Karst.	50.929	13.526	14	1997	2017	28	25	20	8.0	788
Hungary	Puspokszilagy	HCQC	<i>Quercus cerris</i> L.	47.752	19.302	15	1966	2017	77	34	NA	9.2	596
Hungary	Kerecsend	HEQP	<i>Quercus petraea</i> (Matt.) Liebl.	47.815	20.354	15	1977	2015	52	35	NA	9.8	583
Poland	Gleboki Brod	GBPS	<i>Pinus sylvestris</i> L.	53.950	23.210	14	1966	2017	73	40	31	6.9	564
Poland	Jaworzno	JAPS	<i>Pinus sylvestris</i> L.	50.194	19.315	15	1964	2017	82	30	22	8.2	731
Poland	Magdalenga	MAPS	<i>Pinus sylvestris</i> L.	52.080	20.941	15	1970	2017	62	36	22	8.2	542
Poland	Ostrow	OMPS	<i>Pinus sylvestris</i> L.	52.660	21.730	15	1960	2017	92	42	29	7.8	584
Poland	Rogów	ROQR	<i>Quercus robur</i> L.	51.806	19.911	15	1950	2016	86	42	29	8.1	585
Romania	Dumbravita	DBQR	<i>Quercus robur</i> L.	45.769	25.478	15	1903	2017	132	52	27	7.6	716
Romania	Reci	REPS	<i>Pinus sylvestris</i> L.	45.816	25.943	14	1959	2016	107	50	34	7.8	665
Russia	Raifa	RAPS	<i>Pinus sylvestris</i> L.	55.909	48.733	14	1900	1981	202	NA	20	3.9	527
Serbia	Avala	AVPN	<i>Pinus nigra</i> J.F.Arnold	44.672	20.541	12	1977	2017	49	34	20	12.0	642
Serbia	Deliblatski Pesak	DPPS	<i>Pinus sylvestris</i> L.	44.811	21.239	14	1976	2017	50	27	18	11.7	641
Serbia	Stepin Lug	SLPN	<i>Pinus nigra</i> J.F.Arnold	44.748	20.531	15	1976	2017	53	28	18	12.0	642
Switzerland	Albisboden	ABPA	<i>Picea abies</i> (L.) H.Karst.	47.269	8.525	15	1990	2017	37	43	27	8.3	1435
United Kingdom	North England	UKPA	<i>Picea abies</i> (L.) H.Karst.	54.000	-2.400	14	1998	2017	25	18	14	8.6	1273

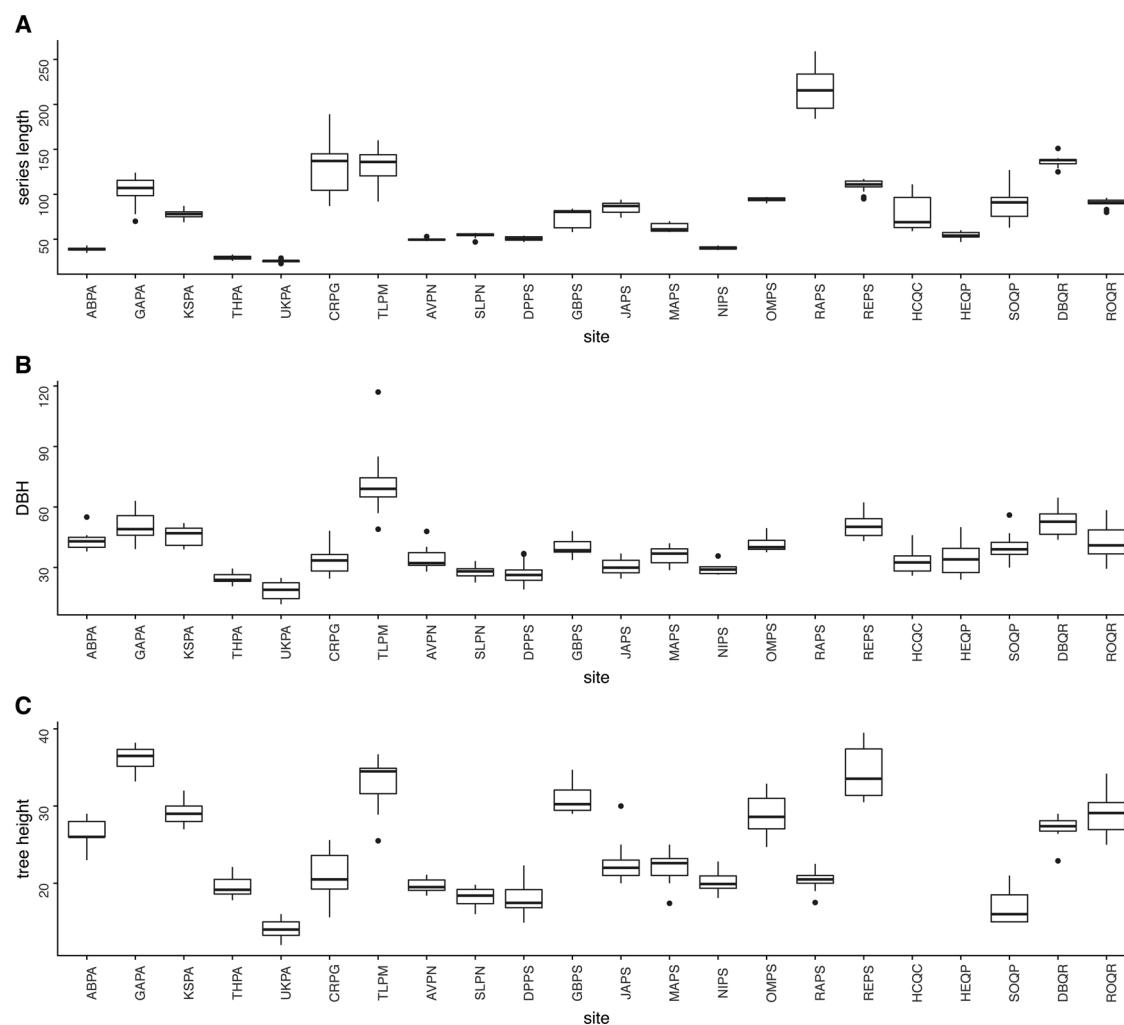


Fig. 1. Descriptive statistics of the studied sites. A: Length of ring-width series in years (per tree only the longer series among the eastern and southern core series was included). B: Diameter at breast height (DBH) in cm. C: Tree height in m. Sites are sorted by species. The last two letters of the site abbreviations provide the species abbreviations (see Table 1).

2.3. Principal Component Gradient Analysis

The *Principal Component Gradient Analysis* (PCGA) was proposed by Buras et al. (2016) as a means of detecting subpopulation growth-signals in tree-ring time-series populations. The method is based on the *Principal Component Analysis* (PCA, Mardia et al., 1979; Venables and Ripley, 2002) but focuses on interpreting the polar-transformation of the loadings of the first two principal components (PC1 and PC2) to identify subpopulations of time-series on respective PCGA-plots (Buras et al., 2016).

A PCGA-plot relies on the loadings (arrows) of the typical PCA bi-plot (e.g. Gower and Hand, 1996). Each of these arrows is representable by the polar transformation (i.e. angle and distance to the origin) of the PC1 and PC2 loadings calculated for the respective time-series (Buras et al., 2016). Based on the polar angles, the underlying time-series are ranked, i.e. sorted with respect to one of the two marginal arrows (Buras et al., 2016). Consequently, one of the two extreme arrows is assigned the rank one, whereas the other is ranked last. The remaining time-series are distributed in between in dependence of their differences in polar angles to the series with the rank one. Distinct, visual groupings of vectors indicate subpopulation signals, provided that PC1 and PC2 explain a reasonable ratio of the variance of the total time-series population.

In the context of this study, PCGA was used to assess the possible occurrence of two direction-specific subpopulations of RWI series (i.e.,

eastern and southern cores) at each site. To quantitatively identify significant direction-specific growth differences, the angles of the polar coordinates were tested at each site for significant location shifts (from the non-parametric mean rank) via a Wilcoxon Signed-Rank test (Bauer, 1972; Hollander et al., 2015). A paired test was feasible, because for each individual tree, an eastern and southern core were sampled, i.e. the observations were paired. To account for the multiplicity of tests, the *p*-values were adjusted by controlling the false discovery rate (Benjamini and Hochberg, 1995).

2.4. Climate correlation analysis

Diverging direction-specific climate-growth relationships were studied based on a correlation analysis. This analysis comprised three steps: Firstly, mean site chronologies were computed from the individual, autoregressive residual series to obtain one master chronology per site representative of the corresponding annual average RWI value. Based on these master chronologies, monthly climate correlations (Spearman, see above) were evaluated over a period comprising previous year April to current year October. By investigating climate-correlations to the master chronologies, we ensured that our data set comprised chronologies that are sensitive to climate. If in this step only insignificant climate-correlations were detected, the investigation of direction-specific effects would be rendered obsolete. We considered mean monthly temperature (temp), monthly precipitation sum (prec),

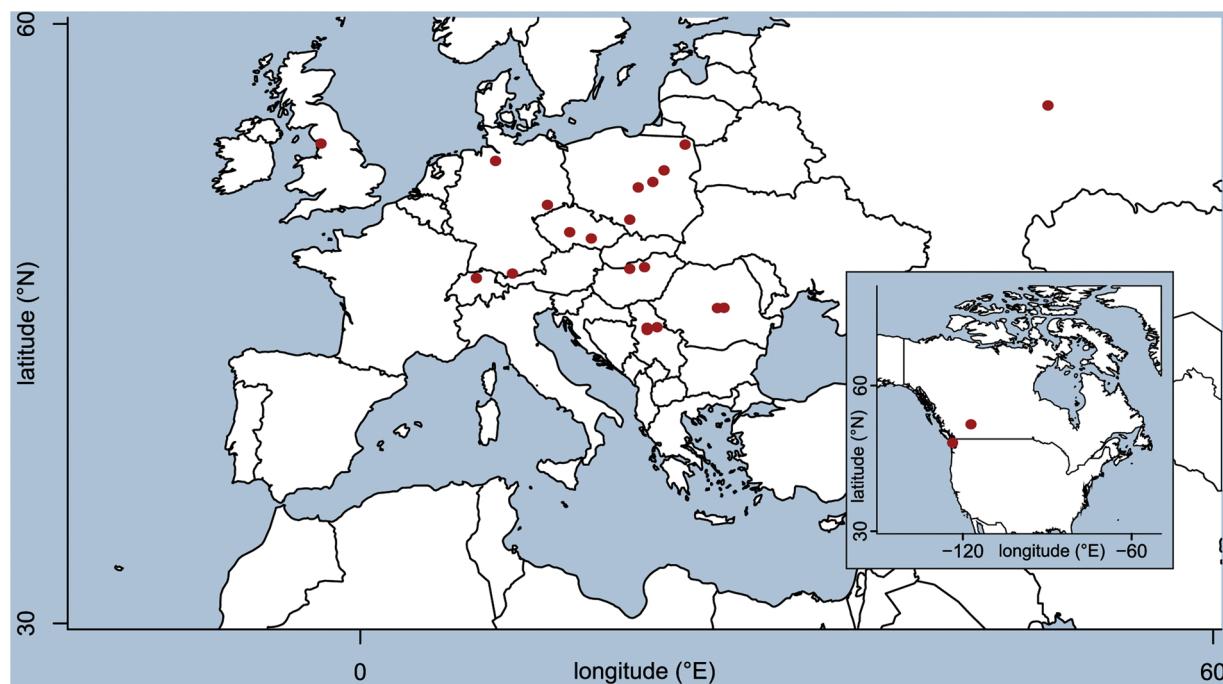


Fig. 2. Geographical location of the study sites (red points) in Europe. Moreover, the study includes two sites located in Canada (inset map). For more meta data see Table 1 (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.).

and the Standardized Precipitation Evapotranspiration Index (SPEI, Vicente-Serrano et al., 2010) as meaningful climate parameters. The SPEI was integrated over three different time-windows: 1, 3 and 6 months (SPEI1, SPE3, SPE6). This accounts for short- and long-term effects of water availability on tree-growth (Vicente-Serrano et al., 2012). For each site, we identified the climate variable(s) with significant correlations to the site chronologies.

Secondly, individual RWI of tree-cores were correlated to the climate variables. The resulting correlation coefficients for respective eastern and southern cores were squared (r^2). The r^2 of a univariate linear regression specifies the variance explained by the regressor (Stock and Watson, 2011). Thus, in our context, r^2 allows for roughly estimating the effect of a given RWI series on a potential reconstruction of the respective climate variable. Hereon, the direction-specific differences in r^2 were investigated via a Wilcoxon Signed-Rank test, which tested the pooled individual differences for eastern and southern cores. Consequently, the test allowed for identifying significant mean

differences between eastern and southern cores' r^2 ($\Delta r^2 = r_{\text{East}}^2 - r_{\text{South}}^2$).

Thirdly, the Δr^2 was calculated for each site and visualized with a heatmap. That is, a positive Δr^2 indicates a higher explained variance of eastern cores, whereas a negative Δr^2 indicates a higher explained variance of southern cores. This difference in Δr^2 allows for roughly estimating the effect of the potential direction-specific bias introduced in a potential climate reconstruction. There was a total of 2090 Wilcoxon Signed-Rank tests on Δr^2 . Hence, the p -values were adjusted by controlling the false discovery rate (Benjamini and Hochberg, 1995).

3. Results

3.1. Differences in signal strength

Generally, direction-specific differences in signal strength were very small and distributed unsystematically (Fig. 3). The mean difference in

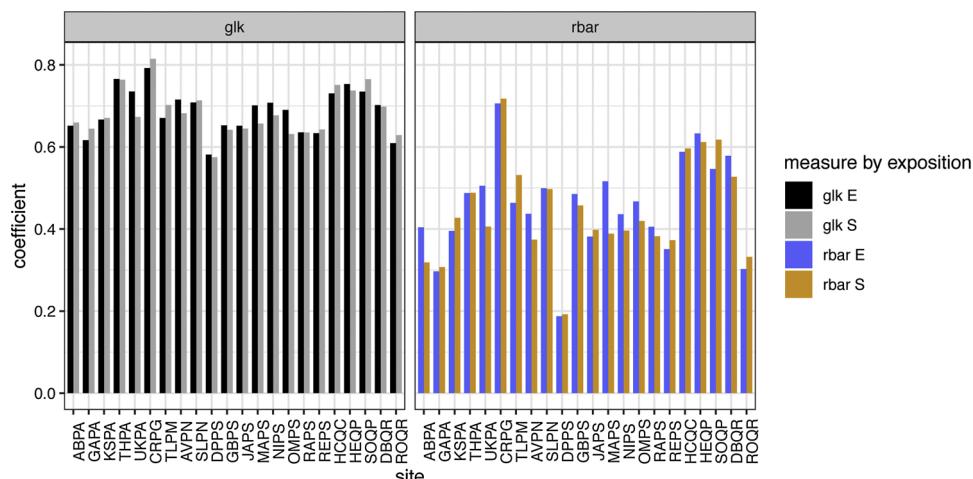


Fig. 3. Direction-specific differences in signal strength between eastern (E) and southern (S) tree cores per site (abbreviations Table 1). Measures are Gleichläufigkeit (glk) and mean inter-series correlation (rbar). Sites are sorted by species.

r between eastern and southern core chronologies (eastern minus southern) was 0.01 (range -0.07 to 0.13). The number of sites that exhibited a lower r for eastern cores was equal to the number of sites with a lower r for southern cores (ratio 0.5). For *glk*, the mean difference was 0.004 (range -0.03 to 0.06). In addition, the number of sites that showed a lower *glk* for eastern cores than for southern cores almost

equaled the number of sites, which showed the opposite (ratio 0.45).

3.2. Direction-specific differences according to PCGA

The ratio of variance explained by PC1 and PC2 was on average 0.56 (range 0.35 to 0.77). The ratio was ≤ 0.5 for only 6 sites (i.e., ABPA, DPPS, GAPA, JAPS, REPS and ROQR), with only one site (DPPS)

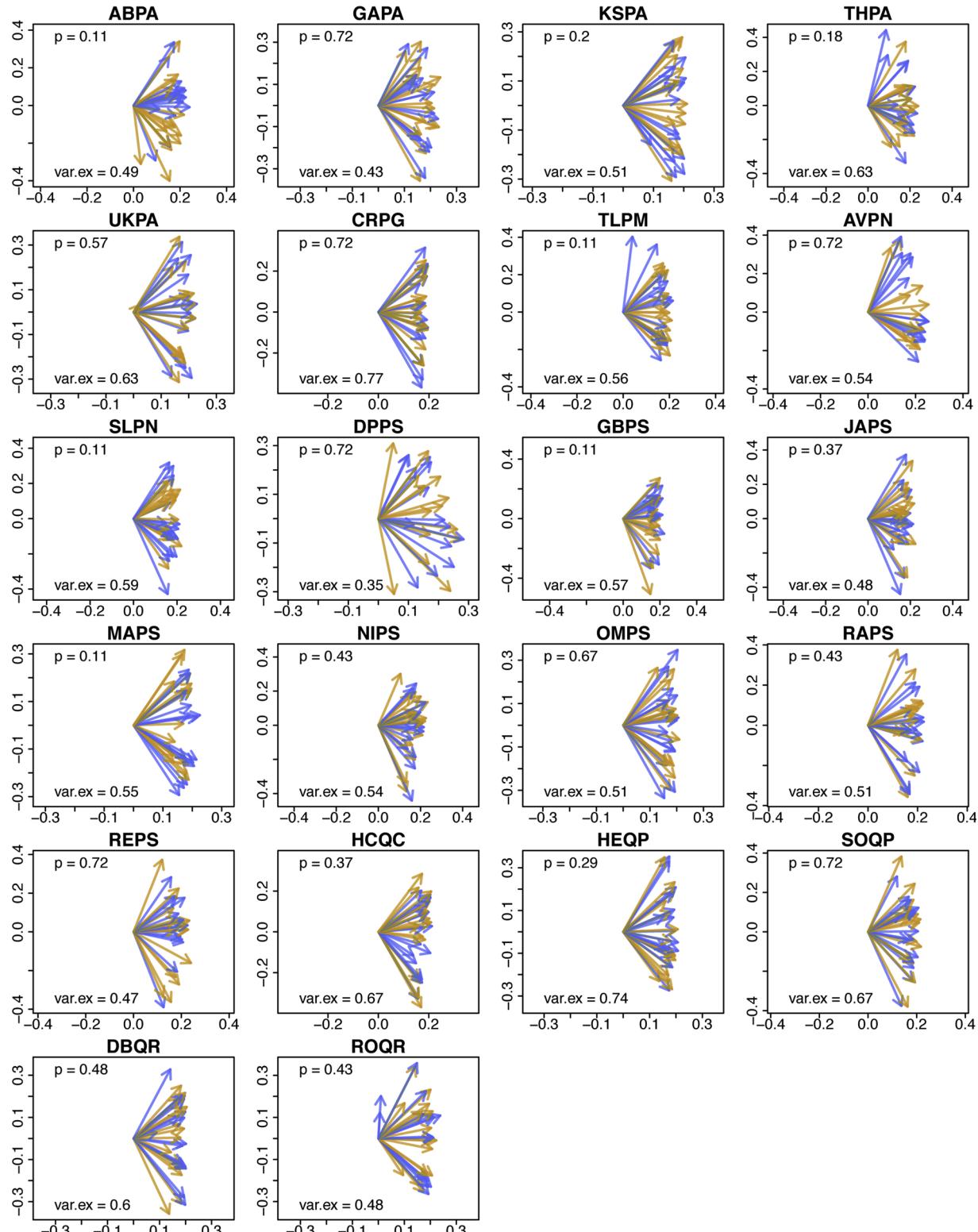


Fig. 4. Direction-specific differences in growth variability according to PCGA. P: p-value of Wilcoxon Signed-Rank tests. Var.ex: Ratio of variance explained by PC1 and PC2. Orange arrows refer to southern, blue arrows to eastern tree cores. Sites are sorted by species (abbreviations Table 1) (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.).

below 0.4 (Fig. 4).

The Wilcoxon Signed-Rank tests detected no significant ($p \leq 0.05$) shifts in the polar angles of the PCGA ranks. Moreover, the PCGA-plots showed no distinct visual groupings of the PCGA ranks (Fig. 4).

3.3. Direction-specific differences in climate-growth correlations

At least one climate variable with a significant ($p \leq 0.05$) correlation to the master chronology was detected per site (Fig. A1). Generally, chronologies were positively correlated with current summer precipitation and current summer SPEI1 to SPEI6, and negatively with summer temperature (Fig. A1).

Direction-specific differences (Δr^2) in climate-growth relationships were insignificant according to the Wilcoxon Signed-Rank tests (Fig. 5). The mean Δr^2 per site were rather small (25% quantile: -0.01; 75% quantile: 0.01; range: -0.06 to 0.09) and showed no systematic distribution (Fig. 5).

4. Discussion

4.1. Multi-national data set shows no evidence of direction-specific tree-ring bias

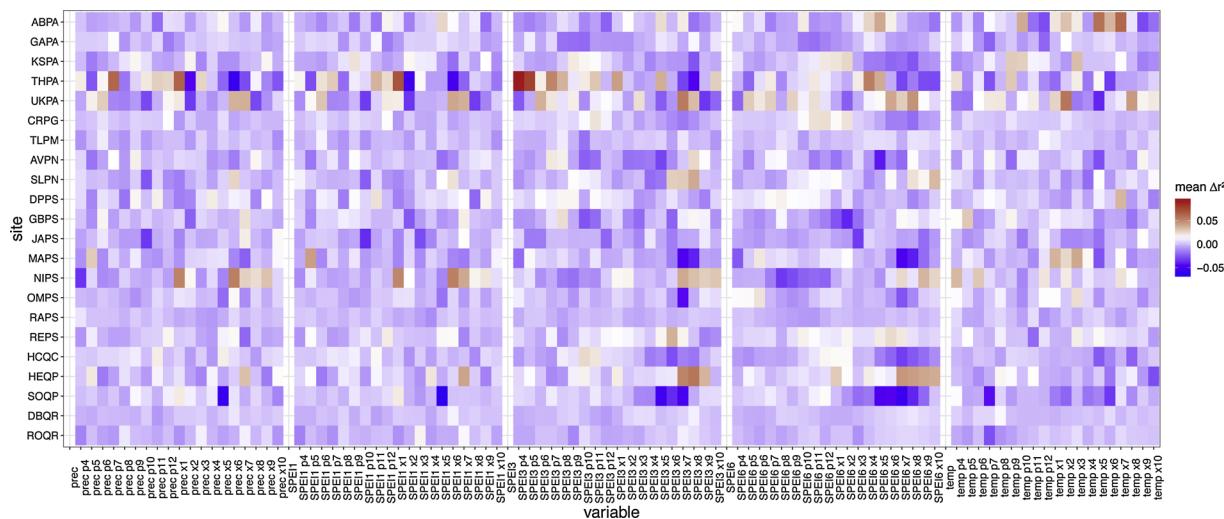
Using a large, multi-species data set with samples derived from sites predominantly representing Europe, we found no evidence for bias due to the cardinal direction of coring. We found virtually no differences between the signal strength statistics for eastern and southern core-chronologies, respectively. Furthermore, the PCGA provided no evidence for direction-specific growth variability. None of the Wilcoxon Signed-Rank tests detected any significant shifts in the polar coordinates. Moreover, the PCGA plots showed no clear visual separation, compared to plots in Buras et al. (2018); Rehschuh et al. (2017). Finally, the absolute differences in climate correlations were marginal (cf. Results), and the Wilcoxon Signed-Rank tests detected no significant, direction-specific differences in Δr^2 . These results corroborate the reliability of many existing tree-ring data sets used in several important fields of research. These results further indicate that conventional sampling protocols (two cores at diameter at breast height) do not introduce any notable direction specific bias in ring-width chronologies. Although many of our sites were mesic and on flat terrain by design (to reduce confounding factors), the large number of sites (22) tested across many countries likely captured a relatively wide range of conditions. Nevertheless, sampling in more extreme environments may show different results.

4.2. Direction-specific growth in previous research

The sites studied by Fang et al. (2015) were dominated by *Pinus tabuliformis* and *Picea purpurea* and mostly located on summits and near cliffs within the semi-arid Chinese Loess Plateau. At such sites, trees are presumably more exposed to environmental factors (e.g. drought stress) than trees at the forest interior (Fritts, 1976; Schweingruber, 1996). Thus, trees might have been more affected by unequal direction-specific exposure to wind and solar radiation than the forest-interior trees studied here. Moreover, if the lack of shade from neighboring trees and the prevailing wind direction are reasonably stable in time, they should result in diverging medium- (3–10 years) to low-frequency (> 10 years) growth signals (Cook and Kairiukstis, 1990). However, these medium- to low-frequency signals are removed by high-pass filtering as used in our study. The high-frequency growth signal is linked to yearly weather conditions (Cook and Kairiukstis, 1990; Fritts, 1976). For direction-specific year-to-year growth variability to consistently occur, climatic factors would need to divergently drive yearly radial tree-growth depending on the cardinal direction over the course of several decades. Thus, while factors like different shading and wind exposition possibly cause direction-specific growth to occur in the mid- to low-frequency, the high-frequency signal seems to be largely unaffected by direction-specific growth variability. Still, in extreme environments, direction-specific effects on the high-frequency growth signal are possible (Buras and Wilmking, 2014; Fang et al., 2015).

For certain shrub species growing in cold tundra ecosystems or forming alpine tree lines, direction-specific growth signals have been reported (Buras and Wilmking, 2014; Shetti et al., 2018; Yang et al., 2018). However, the latter signals were not related to cardinal directions (i.e. geographic exposition). These deviating signals were most likely caused by micro-environmental conditions (Buras and Wilmking, 2014), which may also have been a cause for direction-specific growth in the summit and near-cliff stands studied by Fang et al. (2015).

Consequently, direction-specific growth could occur in extreme environments. Previous studies often focused on diverging growth along the stem rather than on direction-specific growth. For example, Chhin et al. (2010) found changing climatic responses in stem disk mean chronologies taken at different heights and Gričar et al. (2006) showed that heating and cooling of stem sections can affect cambial activity and cell differentiation in *Picea abies*. Thus, research focusing on extremely dry or cold environments could contribute to a more complete assessment of direction-specific growth. In addition, direction-specific climate signals may be investigated in frequencies other than the high-frequency time domain.



4.3. Constraints and limitations of our study

The objective of our study was to evaluate the size of direction-specific climate signals in high-pass filtered RWI series. Instead of stacking several statistical methods, we chose an approach that focuses on the direct assessment of differences in climate correlations and signal strength. We argue that these differences suffice for estimating the relevance of potential direction-specific growth for climate reconstructions. Moreover, this approach is less prone to the problem of interpreting data derivates that result when chaining several multivariate analyses in a sequence, i.e. using the output of an analysis as direct input to the next analysis. Such chaining risks a stepwise disconnection from the original signal differences present in the raw or detrended data, which in turn can lead to a loss of interpretability of the results. Nonetheless, there are several limitations to our study.

4.3.1. Analysis related limitations

For certain sites (i.e., ABPA, DPPS, GAPA, JAPS, REPS and ROQR) the PC1 and PC2 explained less than 50% of the variance. Therefore, additional PCs may have also held meaningful information. However, each of these additional PCs explained an even lower percentage of variance. Moreover, the implementation of polar coordinates in three dimensions is not easily solved. Thus, we restricted our analyses to PC1 and PC2, which explained a relatively high portion of variance for most of the sites.

The PCGA plots showed no clear visual separation. The two-colored PCGA plot is suboptimal for visualizing the Wilcoxon Signed-Rank test, which tests for significant shifts in angles between the pairs of PCGA vectors from the same tree. If the vector pairs have a systematic order, i.e. the eastern is always above or below the southern in the fan of vectors, respectively, the paired test detects significant shifts. It is impossible to visualize these shifts with only two colors, one for each group of eastern and southern cores. Plots with colored vectors specific to the individual tree would overcome this limitation. In addition, a specific line type could be used to indicate if a vector represents a southern or eastern core. However, such plots quickly become fairly complex and may confuse observers (Fig. A4). Thus, and because no significant shifts in polar coordinates were detected, such plots are not shown here.

We used gridded climate data (0.5°) as data from nearby meteorological stations was available just for a minority of sites. The gridded data sets only approximately represent the actual climatic conditions at the sites. Thus, the respective climate-growth relationships could have been underestimated for sites that lie in areas with high micro-climatic variability.

4.3.2. Material/sampling related limitations

While there was not always equal representation of species across sites, there is evidence that this does not bias our results. Some species were represented by several sites with differing ecological conditions (e.g. *Pinus sylvestris* L.), whereas for other species we have studied only one site (e.g. *Pseudotsuga menziesii* (Mirb.) Franco). Yet, the well-represented species (e.g. *Pinus sylvestris* L.) showed no clear evidence for systematic direction-specific growth variability. In addition, the geographic and ecological gradients covered by our study were vast, albeit biased towards Europe. Bearing the two Canadian sites in mind, we speculate that a wider spatial coverage (i.e., the sampling of more sites) across the Northern Hemisphere would probably not reveal any direction-specific biases for relatively mesic sites within flat terrain forests.

We only investigated two cores per tree. Nevertheless, the two major factors that potentially drive direction-specific climate signals in our study area, i.e. solar radiation from the south and wind from the west (for European sites in flat terrain mainly), should have been detectable by coring the southern and eastern sides of the stem only.

As stated above, we acknowledge that direction-specific growth variability and climate signals might occur in extreme environments. Here, we concentrated on flat terrain forest sites. On slopes, direction-specific growth is influenced by reaction wood, which is a factor that is potentially difficult to quantify. The formation of reaction wood is

thought to primarily influence the mid- to low-frequency growth signal (Duncker and Spiecker, 2008). Still, the high-frequency might be affected by reaction wood formation, especially if the physical strain on a stem changes abruptly. In such cases, the climatic effects and the effects related to reaction wood formation are difficult to distinguish and quantify separately (Speer, 2010). Thus, we focused on sites in flat terrain to test the hypothesis of direction-specific climate signals in the high-frequency growth signal of ring-width series.

Finally, the majority of sites represent managed, largely mono-specific forests. It remains to be tested using other data sets, whether our findings are representative for naturally-grown, multi-species forests. However, the two naturally-grown sites in Canada with few individuals of co-dominant tree species revealed similar results as managed European forests. Thus, these results suggest that direction-specific climate signals probably are also negligible for naturally-grown sites similar to the two Canadian sites.

5. Conclusions

This study found no evidence for systematic direction-specific radial growth in high-pass filtered ring-width series. The climate correlations investigated were largely unaffected by direction-specific effects. Hence, the potential for direction-specific bias in ring-width data sets used by the diverse fields of tree-ring research (e.g. the establishment of reference chronologies for dendro-provenancing and millennia-long chronologies for climate reconstructions) seems negligible. At least this appears to be the case for sites that feature similar site conditions to those studied here, i.e. that lie at the forest interior on comparably moderate sites under relatively mesic conditions. Nevertheless, the cardinal direction of core extraction should be documented to control for micro-environmental factors such as the formation of reaction wood due to exposure to a prevailing wind-direction and/or slope inclination.

Direction-specific growth could prove to be more pronounced at extreme (cold or dry) sites, where trees and shrubs are more exposed to severe weather conditions. Thus, the possible consequences of direction-specific growth on dendrochronological studies in extreme environments would make an excellent avenue for future research to evaluate methodologies and data sets at species margins.

Data availability

The tree-ring and meta data sets are available upon request.

Author contribution statement

The scientific concept and research question of this study was developed in collaboration by A. Buras and U. Gut. The latter also conducted the bulk of the analyses presented here. A. Buras initiated the tree-ring data network. Both together have worked on the submitted manuscript. All other members of the network are listed alphabetically as co-authors, with the exception of Ute Sass-Klaassen, who hosted the first network workshop at Wageningen University and Research. These co-authors contributed to the article by providing tree-ring data sets and valuable feedback at various stages of data analyses as well as commenting on the final manuscript version before submission. We are very grateful for the financial support provided by the Association of Tree-ring Research (ATR), which allowed the implementation of a network-internal statistical workshop.

Acknowledgements

U. Gut thanks N. Bleicher for comments and discussions. U. Gut was financially supported by the Swiss National Science Foundation, grant no. P0ZHP1_162299. A. Buras received funding from the German Academic Exchange Service (DAAD). J. J. Camarero acknowledges the support of the CGL2015-69186-C2-1-R project (Spanish Ministry of Economy). R. Cruz-

García was supported by a DAAD-Conacyt scholarship. A. Hevia was supported by OLDPINE (AGL2017-83828-C2-2R) project (Spanish Ministry of Economy, Industry and Competitiveness, MINECO). M. Isaac-Renton and D. Montwé thank Anne-Marie Marchi, Todd Columbia and CRD Regional parks for sampling permission. T. Kolář and M. Rybníček were supported by the Ministry of Education, Youth, and Sports of CR within the National Sustainability Program I (NPUI), grant number LO1415, the Czech Science Foundation (18-17295S). A.M. Petritan was supported by the Ministry of Research and Innovation, CNCS – UEFISCDI, project number PN-III-P1-1.1-TE-2016-1508, within PNCDI III (BIOCARB). R. Sánchez-Salguero was supported by the projects CoMo-ReAdapt (CGL2013-48843-C2-1-R, Spanish Ministry of Economy, Industry and Competitiveness, Spain) and LESENS (RTI2018-096884-B-C33, Ministry of Science, Innovation and Universities, Spain). We thank the two anonymous reviewers for their constructive comments and suggestions, which further improved our manuscript.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.dendro.2019.125624>.

References

- Amoroso, M.M., Daniels, L.D., Baker, P.J., Camarero, J.J. (Eds.), 2017. *Dendroecology: Tree-Ring Analyses Applied to Ecological Studies*. Springer, Cham 400 pp.
- Babst, F., Bodesheim, P., Charney, N., Friend, A.D., Girardin, M.P., Klesse, S., Moore, D.J.P., Seftigen, K., Björklund, J., Bouriaud, O., Dawson, A., DeRose, R.J., Dietze, M.C., Eckes, A.H., Enquist, B., Frank, D.C., Mahecha, M.D., Poultre, B., Record, S., Trouet, V., Turton, R.H., Zhang, Z., Evans, M.E.K., 2018. When tree rings go global: challenges and opportunities for retro- and prospective insight. *Quat. Sci. Rev.* 197, 1–20. <https://doi.org/10.1016/j.quascirev.2018.07.009>.
- Bauer, D.F., 1972. Constructing confidence sets using rank statistics. *J. Am. Stat. Assoc.* 67 (339), 687–690. <https://doi.org/10.1080/01621459.1972.10481279>.
- Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 57 (1), 289–300.
- Best, D.J., Roberts, D.E., 1975. Algorithm AS 89: the upper tail probabilities of Spearman's rho. *J. R. Stat. Soc. Ser. C Stat. Methodol.* 24 (3), 377. <https://doi.org/10.2307/2347111>.
- Bleicher, N., Harb, C. (Eds.), 2015. *Zürich-Parkhaus Opéra: Eine neolithische Feuchtbodenfundstelle. FO-Fototorar*, Zürich, Egg 270 pp.
- Bridge, M., 2012. Locating the origins of wood resources: a review of dendroprovenancing. *J. Archaeol. Sci.* 39 (8), 2828–2834. <https://doi.org/10.1016/j.jas.2012.04.028>.
- Bunn, A.G., 2008. A dendrochronology program library in R (dplR). *Dendrochronologia* 26 (2), 115–124. <https://doi.org/10.1016/j.dendro.2008.01.002>.
- Bunn, A., Korpela, M., Biondi, F., Campelo, F., Mérian, P., Qeadan, F., Zang, C., Puchacofrep, D., Wernicke, J., 2018. dplR: Dendrochronology Program Library in R.
- Buras, A., Schunk, C., Zeiträg, C., Herrmann, C., Kaiser, L., Lemme, H., Straub, C., Taeger, S., Gößwein, S., Klemmt, H.-J., Menzel, A., 2018. Are Scots pine forest edges particularly prone to drought-induced mortality? *Environ. Res. Lett.* 13 (2), 25001. <https://doi.org/10.1088/1748-9326/aaa0b4>.
- Buras, A., van der Maaten-Theunissen, M., van der Maaten, E., Ahlgrimm, S., Hermann, P., Simard, S., Heinrich, I., Helle, G., Unterseher, M., Schnittler, M., Eusemann, P., Wilming, M., 2016. Tuning the voices of a choir: detecting ecological gradients in time-series populations. *PLoS One* 11 (7), e0158346. <https://doi.org/10.1371/journal.pone.0158346>.
- Buras, A., Wilming, M., 2014. Straight lines or eccentric eggs? A comparison of radial and spatial ring width measurements and its implications for climate transfer functions. *Dendrochronologia* 32 (4), 313–326. <https://doi.org/10.1016/j.dendro.2014.07.002>.
- Buras, A., Wilming, M., 2015. Correcting the calculation of Gleichläufigkeit. *Dendrochronologia* 34, 29–30. <https://doi.org/10.1016/j.dendro.2015.03.003>.
- Carrer, M., 2011. Individualistic and time-varying tree-ring growth to climate sensitivity. *PLoS One* 6 (7), e22813. <https://doi.org/10.1371/journal.pone.0022813>.
- Chhin, S., Hogg, E.H.T., Lieffers, V.J., Huang, S., 2010. Growth-climate relationships vary with height along the stem in lodgepole pine. *Trees Physiol.* 30 (3), 335–345. <https://doi.org/10.1093/treephys/tpp120>.
- Cook, E.R., Kairikstis, L.A., 1990. *Methods of Dendrochronology*. Springer, Netherlands, Dordrecht.
- Duncker, P., Spiecker, H., 2008. Cross-sectional compression wood distribution and its relation to eccentric radial growth in *Picea abies* [L.] Karst. *Dendrochronologia* 26 (3), 195–202. <https://doi.org/10.1016/j.dendro.2008.06.004>.
- Eckstein, D., Bauch, J., 1969. Beitrag zur Rationalisierung eines dendrochronologischen Verfahrens und zur Analyse seiner Aussagesicherheit. *Forstwissenschaftliches Centralblatt* 88, 230–250.
- Eissing, T., Dittmar, C., 2011. Timber transport and dendroprovenancing in Thuringia and Bavaria. *Treerings, Art and Archaeology*. pp. 137–149.
- Fang, K., Chen, D., Gou, X., D'Arrigo, R., Davi, N., 2015. Influence of non-climatic factors on the relationships between tree growth and climate over the Chinese Loess Plateau. *Glob. Planet. Change* 132, 54–63. <https://doi.org/10.1016/j.gloplacha.2015.06.008>.
- Fritts, H.C., 1976. *Tree Rings and Climate*. Elsevier.
- Galván, J.D., Camarero, J.J., Gutiérrez, E., 2014. Seeing the trees for the forest: drivers of individual growth responses to climate in *Pinus uncinata* mountain forests. *J. Ecol.* 102 (5), 1244–1257. <https://doi.org/10.1111/1365-2745.12268>.
- Gower, J.C., Hand, D.J., 1996. *Biplots*, 1st ed. Chapman & Hall, London 277 pp.
- Gričar, J., Zupancic, M., Cufar, K., Koch, G., Schmitt, U., Oven, P., 2006. Effect of local heating and cooling on cambial activity and cell differentiation in the stem of Norway spruce (*Picea abies*). *Ann. Bot.* 97 (6), 943–951. <https://doi.org/10.1093/aob/mcl050>.
- Grissino-Mayer, H., 2001. *Evaluating crossdating accuracy: a manual and tutorial for the computer program COFECHA*. Tree-Ring Res. 57.
- Gut, U., 2018. Evaluating the key assumptions underlying dendro-provenancing: how to spruce it up with a scissor plot. *Dendrochronologia* 52, 131–145. <https://doi.org/10.1016/j.dendro.2018.09.008>.
- Harris, I., Jones, P.D., Osborn, T.J., Lister, D.H., 2014. Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 dataset. *Int. J. Climatol.* 34 (3), 623–642. <https://doi.org/10.1002/joc.3711>.
- Hollander, M.A., Wolfe, D., Chicken, E., 2015. *Nonparametric Statistical Methods*. John Wiley & Sons, Inc, Hoboken, NJ, USA.
- Hughes, M.K., Swetnam, T.W., Diaz, H.F., 2011. *Dendroclimatology*. Springer, Netherlands, Dordrecht.
- Klesse, S., DeRose, R.J., Guiterman, C.H., Lynch, A.M., O'Connor, C.D., Shaw, J.D., Evans, M.E.K., 2018. Sampling bias overestimates climate change impacts on forest growth in the southwestern United States. *Nat. Commun.* 9 (1), 5336. <https://doi.org/10.1038/s41467-018-07800-y>.
- Mardia, K.V., Kent, J.T., Bibby, J.M., 1979. *Multivariate Analysis*. Acad. Press, London 521 pp.
- Melvin, T.M., Briffa, K.R., 2008. A "signal-free" approach to dendroclimatic standardization. *Dendrochronologia* 26 (2), 71–86. <https://doi.org/10.1016/j.dendro.2007.12.001>.
- PAGES 2k Consortium, 2013. Continental-scale temperature variability during the past two millennia. *Nat. Geosci.* 6 (5), 339–346. <https://doi.org/10.1038/ngeo1797>.
- Peters, R.L., Groenendijk, P., Vlam, M., Zuidema, P.A., 2015. Detecting long-term growth trends using tree rings: a critical evaluation of methods. *Glob. Change Biol.* 21 (5), 2040–2054. <https://doi.org/10.1111/gcb.12826>.
- R Core Team, 2018. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rehschuh, R., Mette, T., Menzel, A., Buras, A., 2017. Soil properties affect the drought susceptibility of Norway spruce. *Dendrochronologia* 45, 81–89. <https://doi.org/10.1016/j.dendro.2017.07.003>.
- Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., Ziese, M., 2016. GPCC Full Data Reanalysis Version 7.0: Monthly Land-surface Precipitation From Rain Gauges Built on GTS Based and Historic Data.
- Schweingruber, F.H., 1996. *Tree Rings and Environment Dendroecology*. Haupt, Berne, Stuttgart, Vienna 609 pp.
- Shetti, R., Buras, A., Smiljanic, M., Wilmking, M., 2018. Climate sensitivity is affected by growth differentiation along the length of *Juniperus communis* L. shrub stems in the Ural Mountains. *Dendrochronologia* 49, 29–35. <https://doi.org/10.1016/j.dendro.2018.02.006>.
- Speer, J.H., 2010. *Fundamentals of Tree-Ring Research*. Univ. of Arizona Press, Tucson, Ariz 333 pp.
- Stock, J.H., Watson, M.W., 2011. *Introduction to Econometrics*, 3rd ed. Pearson/Addison Wesley, Boston XLII, 785.
- Sullivan, P.F., Csank, A.Z., 2016. Contrasting sampling designs among archived datasets: implications for synthesis efforts. *Tree Physiol.* 36 (9), 1057–1059. <https://doi.org/10.1093/treephys/tpw067>.
- Venables, W.N., Ripley, B.D., 2002. *Modern Applied Statistics With S*. Springer, New York, NY 497 pp.
- Vicente-Serrano, S.M., Beguería, S., López-Moreno, J.I., 2010. A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. *J. Climate* 23 (7), 1696–1718. <https://doi.org/10.1175/2009JCLI2909.1>.
- Vicente-Serrano, S.M., Beguería, S., Lorenzo-Lacruz, J., Camarero, J.J., López-Moreno, J.I., Azorin-Molina, C., Revuelto, J., Morán-Tejeda, E., Sanchez-Lorenzo, A., 2012. Performance of drought indices for ecological, agricultural, and hydrological applications. *Earth Interact.* 16 (10), 1–27. <https://doi.org/10.1175/2012EI000434.1>.
- Wigley, T.M.L., Briffa, K.R., Jones, P.D., 1984. On the average value of correlated time series, with applications in dendroclimatology and hydrometeorology. *J. Climate Appl. Meteorol.* 23 (2), 201–213. [https://doi.org/10.1175/1520-0450\(1984\)023<0201:OTAVOC>2.0.CO;2](https://doi.org/10.1175/1520-0450(1984)023<0201:OTAVOC>2.0.CO;2).
- Yang, J., Zhao, H., Zhang, Y., Li, Z., Wang, X., 2018. Climate-growth relationship for different directions of *Pinus pumila* radial growth at the treeline of northern Daxing'an Mountains, China. *Trees* 32 (1), 311–322. <https://doi.org/10.1007/s00468-017-1633-4>.

(3) Artikel III

Assessing site signal preservation in reference chronologies used for dendro-provenancing.

Urs Gut

Dieser Artikel wurde eingereicht bei:

PLOS ONE

1 Assessing site signal preservation in reference
2 chronologies for dendro-provenancing

3 Urs Gut^{1,2*}

4 ¹Fachbereich Prähistorische Archäologie, Institut für Archäologie, University of
5 Zurich, Zurich, Switzerland

6 ²Forest Ecology, Institute of Terrestrial Ecosystems, Department of Environmental
7 System Science, ETH Zurich, Zurich, Switzerland

8 *Corresponding author

9 E-mail: u.gut@posteo.de

10

11

12

13

14

15

16

17

18

19

20

21 1 Abstract

22 Regional differences in tree growth can be used to approximate the geographical
23 provenance of ring-width series ('dendro-provenancing'). This method relies on
24 cross-dated ring-width series (reference chronologies) that are thought to represent
25 the radial growth signal of trees in a given region. Reference chronologies are often
26 established from ring-width series of living tree populations. Frequently, they are too
27 short to allow for investigating the provenance of historical wood. Thus, references
28 are extended by ring-width series from buildings and objects of art that exhibit best
29 matching growth patterns with the living tree references. Yet, series from other
30 provenances may erroneously be included. Thereby the local or regional growth
31 signal of the references is progressively contaminated, but this has received little
32 attention to date.

33 I investigate this contamination risk using a simulation approach that allows for
34 generating pseudo site chronologies that preserve the relevant statistical properties
35 of the real site chronologies. While the exact provenance of historical wood is
36 unknown, for simulated ring-width series the provenance is unambiguous. Hence,
37 pseudo reference chronologies may be established while monitoring the signal
38 mixture. Specifically, 15 site chronologies of Norway spruce (*Picea abies* (L.) H.
39 Karst.) from northeastern Switzerland were used to generate 15 pseudo site growth
40 signals that span 1000 years.

41 The simulation demonstrates that quasi uncontaminated references can be
42 established in ideal circumstances for the study area. However, the thresholds for the
43 similarity in between-series correlation must be very high. Even then, contaminated
44 pseudo references occurred in rare cases during the simulation. Yet, elevation-
45 specific pseudo references were established with lower thresholds. Simulation

46 currently offers the only approach for assessing the contamination risk of reference
47 chronologies, and it allows for elucidating the conditions under which acceptable
48 levels of contamination can be guaranteed. Therefore, the present approach paves
49 the way towards a practical simulation tool for dendro-provenancing.

50 **2 Introduction**

51 Dendrochronology (most commonly, the study of tree-ring width) is used in many
52 scientific disciplines, e.g. in Forestry, Climatology and Ecology [1]. In Archaeology
53 and History, it is among the most important methods for establishing absolute
54 chronological frameworks and providing environmental information [2–4].

55 Dendro-provenancing goes beyond the dating of wooden objects and capitalizes on
56 regional differences in tree growth that originate from regional differences in climatic
57 and site conditions. If such differences are pronounced, dendro-provenancing can be
58 used to approximate the site of growth of trees that were used to produce wooden
59 objects and artifacts [5–8]. Thus, for example, timber trade routes and areas of
60 timber supply of cities can be reconstructed [9,10], or the origin of wooden resources
61 used e.g. for panel painting and shipbuilding can be determined [11–13].

62 The potentials and limitations of dendro-provenancing have been reviewed [14,15],
63 and the key assumptions underlying dendro-provenancing have been formulated and
64 evaluated as well [15]. In short, dendro-provenancing is based on the comparison of
65 candidate ring-width series from historical objects of unknown provenance to a
66 framework of local reference chronologies. The reference chronologies are
67 established from cross-dated samples of ring-width measurements that are thought
68 to represent local tree-growth. The location of the provenance is then approximated

69 by the geographical area in which best matches between the ring-width patterns of a
70 candidate series and the reference chronologies are found [16]. Generally, best
71 matches are determined statistically using the same methods as for the (cross-
72)dating of ring-width series [17,18].

73 In a previous study [15], ring-width chronologies were established from living spruce
74 (*Picea abies* (L.) H. Karst.) at 15 sites in northeastern Switzerland. The respective
75 dataset provides the basis for reference chronologies for dendro-provenancing.
76 Specifically, the living tree chronologies can be extended by best matching historical
77 ring-width series. Indeed, the dataset showed sufficient between-site year-to-year
78 (high-frequency) ring-width variability to potentially allow for dendro-provenancing in
79 the study area [15]. However, the extent to which these differences are preserved
80 when living tree chronologies are extended by successively adding best matching
81 historical ring-width series has received little attention to date [18–21]. For dating, the
82 merging of site growth signals is relatively inconsequential, because the supra-
83 regional climatically determined growth signal is preserved [22]. However, for dendro-
84 provenancing to work, reference chronologies must preserve local to regional growth
85 signals.

86 Unfortunately, the preservation of local and regional growth signals cannot be studied
87 based on historical ring-width series because their geographical provenance often is
88 unknown. At best, the site of growth may be approximated by watersheds [9]. This is
89 applicable to the pre-industrial epoch only, i.e. before the onset of timber trade via
90 railroads. Even earlier, rafting of timber sometimes covered long distances [10,23].
91 Thus, only in areas where past timber trade was spatially limited there is a chance for
92 reference chronologies to capture local growth signals [24]. However, the
93 establishment of such chronologies is extremely time- and labor-intensive. In

94 addition, if the between-site signal differences are unpronounced, the different local
95 references probably merge into a single supra-regional chronology. Hence, work
96 should be focused on promising study regions. Therefore, the potential for
97 establishing reference chronologies for dendro-provenancing needs to be evaluated
98 prior to field work.

99 The evaluation of signal preservation in reference chronologies can be approached
100 via simulation. The basic concept underlying this paper is twofold: Firstly, the most
101 relevant statistical properties of recently (i.e., in 2015 and 2016) collected site
102 chronologies are characterized and an adequate simulation model is developed.
103 Secondly, this model is used to generate a set of simulated historic pseudo ring-width
104 series. The latter are samples from an underlying set of pseudo site chronology
105 signals that preserve the covariance of the original site chronologies. Thirdly, once
106 the pseudo historical tree-ring data have been generated, an algorithm is used that
107 mimics the extension of site chronologies by adding best matching pseudo historical
108 series. Finally, the composition of the generated pseudo reference chronologies is
109 evaluated. The critical point is that the site provenance is known for all pseudo
110 historical series, as they were sampled from the respective underlying pseudo site
111 signals. Hence, it is possible to examine the composition of the pseudo references
112 and determine the degree of signal mixture. The current case study is limited to the
113 material gathered by [15], but the approach is transferable to other dendro-
114 provenancing studies.

115 The objectives of this study relate to the simulation model, the classification algorithm
116 and the evaluation of the algorithmically generated pseudo reference chronologies,
117 and are as follows:

- 118 1. To introduce and evaluate a simulation model for generating pseudo historical
119 ring-width chronologies.
- 120 2. To develop an algorithm that mimics the process of establishing reference
121 chronologies for dendro-provenancing.
- 122 3. To examine the signal composition of the pseudo reference chronologies and
123 assess the potential for establishing real reference chronologies in the study
124 area.

125 **3 Materials and methods**

126 **3.1 Real ring-width dataset**

127 Real ring-width datasets are needed to initialize the simulation. The dataset used for
128 this purpose comprises 401 ring-width series that represent a set of 15 site
129 chronologies of Norway spruce (*Picea abies* (L.) H. Karst.) with a replication of 15 to
130 32 ring-width series per site [15].

131 Further details on the sampling design, measurement, cross-dating and meta data
132 are given in [15]. The ring-width dataset is provided in the Supporting Information (R
133 package data *chrono.rwl* in S1).

134 For the simulation, the actual site name is unimportant. Hence, only abbreviations
135 are used to label the pseudo site signals. However, the letters preceding the dot in
136 the abbreviations (e.g. as used in subsection ‘Pref composition’) are identical to the
137 real site abbreviations introduced in [15]. Thus, the pseudo sites signals can be
138 traced back to their underlying real site signals for more detailed scrutiny.

139 **3.2 Simulation**

140 All calculations and statistical analyses were done in R (3.6.0) [25]. Specifically, an
141 *ad hoc* R package was written to perform the simulation. The complete code can be
142 found in the Supporting Information (S1).

143 Auto-regressive (henceforth: ar) residual series effectively preserve the signal that is
144 relevant for dendro-provenancing [15]. Thus, ar residual chronologies were used for
145 the simulation. Unlike in raw ring-width series, in ar residuals the medium- (5-15
146 years) and low-frequency (>15 years) growth variability is quasi absent [26]. The
147 individual ar residual series hence represent the year-to-year (high-frequency) growth
148 variability of each individual tree ring-width series. Thus, for each year of an ar
149 chronology, there is a sample equal to the number of trees that were sampled and
150 that featured a ring in the respective year (S2 Fig 1A). These yearly distributions of ar
151 residuals may be approximated by normal distributions, as, theoretically, ar residuals
152 are series of normally distributed random shocks [27,28]. Hence, as for any other
153 normally distributed variable, the yearly means and standard deviations of the
154 theoretical population distributions of the ar residuals can be estimated by calculating
155 the sample means and standard deviations of the empirical distributions of the ar
156 residuals [28].

157 Thus, as a first step towards a simulation model of the real tree-ring data, for each of
158 the 15 raw ring-width chronologies an auto-regressive residual chronology was
159 calculated using the default settings of the detrender-function in the R package *dplR*,
160 i.e. by fitting an auto-regressive model to the raw ring-widths and choosing the model
161 that minimized Akaike's Information Criterion [29,30]. This choice of preprocessing
162 was supported by the high classification performance of this method in the study of
163 [15]. When requiring a minimal replication of 15 data points per year of the

164 chronology, the 15 resulting ar chronologies covered a common period of 63 years
165 (i.e., from 1951 to 2014). Subsequently, series of empirical mean values and
166 standard deviations were calculated for these 63 years. This resulted in a 63×30
167 matrix containing 15 column vectors of yearly mean values (rmv) and 15 column
168 vectors of yearly standard deviations (rsd).

169 To sidestep the limited length of the real dataset, a 1000×30 matrix was generated
170 whose columns had zero covariance and followed a standard normal distribution (S2
171 Fig 1B). This matrix was transformed to a matrix with the same scale and covariance
172 as the real data matrix. The transformation matrix was generated by decomposing
173 the covariance matrix of the real rmv-rsd matrix via its eigenvalues [31,32].

174 The R implementation of the function *rmvnorm* (MASS package) assumes a
175 multivariate normal distribution of the underlying variables [32]. Thus, the
176 transformation was valid for covariance matrices calculated from such variables only.
177 For the rmv vectors, a multivariate normal distribution was reasonable to assume
178 because the indices of the ar residuals were themselves normally distributed [27,28].
179 For the rsd vectors, however, a multivariate log-normal distribution seemed more
180 appropriate (S2 Fig 1). Hence, the rsd were log-transformed (natural logarithm) prior
181 to the calculation of the covariance matrix, whose decomposition was used to linearly
182 transform the uncorrelated random data matrix (S2 Fig 1B).

183 After the linear transformation, the generated matrix contained 30 column vectors of
184 length 1000, with 15 columns for the yearly pseudo mean values (pmv) and 15 for
185 the logarithms of the yearly pseudo standard deviations ($\ln(\text{psd})$) of each pseudo
186 chronology. Next, the $\ln(\text{psd})$ were raised to their power to provide the yearly
187 standard deviations of the pseudo chronologies (psd). Finally, the resulting matrix
188 contained all measures of central tendency and dispersion necessary to

189 parameterize the respective sequences of yearly normal distributions for each
190 pseudo site chronology (cf. matrix P , S2 Fig 1B).

191 **3.2.1 Pseudo historic series**

192 To generate individual pseudo series (ps), random samples were sequentially drawn
193 from the sequences of normal distributions defined by the parameters in the pseudo
194 signal matrix. Thus, ps with a length of 1000 years were generated. The ps were cut
195 into partitions (pseudo historic series, phs) that were of similar lengths as the series
196 found in a typical dataset of historical ring-width series. Here, the dataset of the
197 Dendrochronological Laboratory of the City of Zurich was used to determine the
198 partitioning (R package data *distr.s1* in S1).

199 To partition a 1000 years ps, series lengths were drawn randomly from the pool of
200 series lengths of *Picea abies* (length ≥ 50 years) until the sum of the drawn series
201 lengths was ≥ 1000 years. Then the 1000 years ps was cut into subseries. Often, this
202 resulted in the last partition being shorter than 50 years, which were discarded during
203 later analyses (cf. section PREF-Constructor Algorithm, below). The sampling of ps
204 and their subsequent partitioning into phs was repeated until a replication of 30 phs
205 per year and pseudo chronology was achieved.

206 **3.2.2 Pseudo object chronologies**

207 Generally, object-based chronologies are the core components of reference
208 chronologies for dendro-provenancing rather than single tree-ring series. Such
209 object-based chronologies already represent an aggregate (i.e. the mean) of several
210 ring-width series, which were sampled in historical objects (e.g. the roof beams of a
211 building). If timber sources for the construction of respective objects were spatially
212 limited, object-based chronologies represent mean local tree growth [10]. To allow for
213 a comparison of pseudo reference chronologies established from phs and pseudo

214 reference chronologies generated from pseudo object chronologies (poc/pocs), the
215 above procedure of generating phs was slightly modified, as described below.

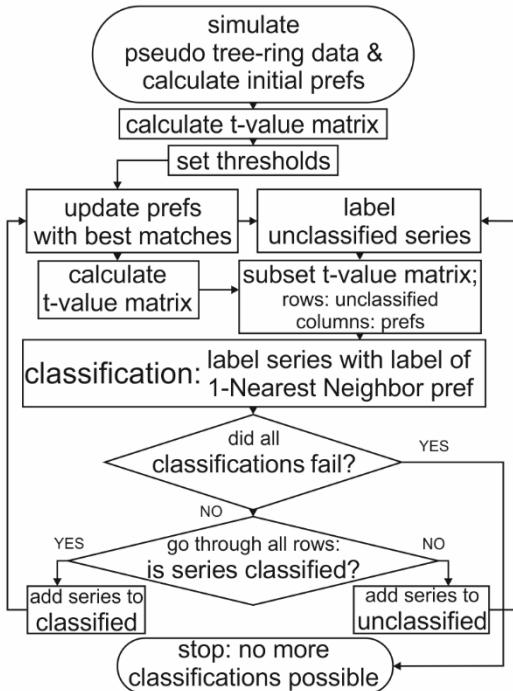
216 Instead of drawing a single ps, 6 ps were sampled from the sequence of yearly
217 normal distributions of a respective pseudo chronology signal. This replication was
218 chosen arbitrarily but seems to be a realistic assumption for the average number of
219 samples taken per historical object in a real setting [33]. Subsequently, the 6 ps were
220 averaged prior to the partitioning (as described in the previous section). Thus, the
221 resulting pocs represent an object chronology with a replication of 6 individual ps per
222 year each. This reflects an ideal scenario, in which the object chronologies are mean
223 value chronologies calculated from phs that originated from the same pseudo site
224 signal. Again, as with the phs approach the sampling and partitioning was repeated
225 until a replication of 30 pocs per year and pseudo chronology was achieved.

226 In a real setting, there is a chance for historical building chronologies to contain a
227 mixture of different site signals [10]. Thus, some simulations were done with pocs
228 containing a mixture of ps originating from different pseudo signals. Specifically, the
229 ratio of ps originating from the correct, ‘local’ pseudo signal was lowered. Thus, this
230 so-called on-site ratio (osr) was lowered from 1 (all 6 ps on-site) to 0.83, i.e. 1 of the
231 6 ps originates from an off-site pseudo signal, to 0.67, i.e. 2 of the 6 ps are off-site
232 ps.

233 **3.3 PREF-Constructor Algorithm**

234 To mimic the establishment of reference chronologies by adding historical ring-width
235 series to initial living tree chronologies and automate this process, a Pseudo
236 Reference Constructor Algorithm (PREF-Constructor) was developed and
237 implemented in the *ad hoc* R package mentioned above (S1). The algorithm

238 generates pseudo reference chronologies (pref/prefs) by adding phs/pocs to the best
 239 fitting initial pref, i.e. to the One-Nearest-Neighbor chronology [34–36]. Thus, prefs
 240 are established that cover 1000 years in ideal circumstances (Fig 1).



241

242 **Fig 1: Flowchart of the PREF-Constructor algorithm.**

243 The 15 initial prefs of this study represent the 15 mean values calculated from 30 ps
 244 of length 150 years per pseudo site chronology. These 30 ps per initial pref were
 245 sampled in addition to the already simulated pseudo dataset. This mimics a scenario
 246 in which a dendrochronologist starts out with 15 unmixed mean site chronologies (the
 247 initial prefs) and extends these by adding the best matching series from the
 248 hypothetical laboratories' historical dataset (here, the simulated phs/poc dataset).

249 Best matches were determined statistically, i.e. by calculating a t-value for each
 250 Pearson's correlation coefficient that resulted from pairwise comparisons between a
 251 candidate series (phs/poc) and the 15 initial prefs [37,38]. This method in
 252 combination with ar prewhitening had a high performance in a previous study [15].

253 Because here the ps resemble ar residual series, a comparable performance was
254 expected.

255 Potential best matches were required to fulfill threshold criteria. Specifically, the
256 thresholds for t-values that were tested were $t \geq 5$ ($t5$), $t \geq 10$ ($t10$), $t \geq 15$ ($t15$) and $t \geq$
257 20 ($t20$) at a minimum overlap of 50 years. If several matches fulfilled the threshold
258 criteria, the respective candidate series was classified to the best matching initial
259 pref, i.e. the chronology that yielded the highest t-value for the pairwise comparison.

260 In a real dataset, potential cross-dating errors complicate the classification. The
261 PREF-Constructor assumes that the cross-dating of each candidate phs/poc is
262 correct. Thus, best matches need to be found for one correlation position only.

263 Once the best matching initial pref had been determined for all phs/pocs, the series
264 were labelled (i.e. classified) according to the label of the best matching initial pref.
265 After this initial run, the first prefs were established from the classified phs/pocs.
266 Subsequently, the phs/pocs that were yet unclassified were compared to the prefs
267 just established. Again, all series that fulfilled the threshold criteria were classified
268 and added to the respective pref. Series that remained unclassified were checked
269 again during later runs, after the establishment of the prefs had advanced. Thus, the
270 algorithm proceeded until all series were classified, or until no unclassified series
271 fulfilled the threshold criteria anymore, i.e. no classifications were possible any more
272 (Fig 1).

273 **3.4 Evaluation**

274 One thousand pseudo datasets comprising 15 pseudo chronologies each were
275 generated to evaluate 1) the simulation model and 2) the PREF-Constructor results.
276 This 1000-fold repetition of the simulation was done for the phs approach as well as

277 for the poc approach and with different settings for the thresholds. In the following
278 two subsections, performance indicators are introduced that provide a basis for
279 quantifying the stability of the pseudo signals across all repetitions. Moreover,
280 methods are presented that allow for the comparison of PREF-Constructor
281 classifications between different simulation settings.

282 **3.4.1 Stability of the pseudo signals**

283 The simulated pseudo signals must remain stable across different simulation
284 repetitions. Instability would indicate that the individual pseudo series (ps) were
285 sampled from a different set of pseudo site signals in each repetition of the
286 simulation. This would render the 1000 pseudo datasets incomparable. Thus, the
287 stability of the intercorrelation between the pseudo mean values (pmv) or the sample
288 pseudo mean value chronologies (spmv), respectively, was investigated.

289 • **$\Delta\{\text{pmv}-\text{rmv}\}$**

290 To quantify the stability of the linear transformation that yielded the pseudo mean
291 values (pmv/pmvs) for the simulation, the pairwise correlations (i.e., the correlation
292 matrix) between the pmv vectors were calculated for each of the 1000 simulation
293 repetitions. Subsequently, the correlation matrix for the mean values of the real site
294 chronologies (rmv/rmvs) was calculated, and the range of differences between the
295 entries of the correlation matrices calculated for the pmvs and for the rmvs ($\Delta\{\text{pmv}-$
296 $\text{rmv}\}$) was determined.

297 • **$\Delta\{\text{spmv}-\text{rmv}\}$**

298 The linear transformation that yielded the pmvs was expected to be stable, except for
299 rounding errors [31]. However, more instability was anticipated for the mean values
300 calculated from phs chronologies. Even if the PREF-Constructor algorithm were to

301 classify each phs correctly, the pseudo reference chronologies (pref) calculated from
302 these phs would not perfectly reproduce the intercorrelation of the pmv. This was to
303 be expected because the phs are random realizations of the respective pseudo
304 signals (section Simulation). Thus, the variability introduced due to random sampling
305 needed to be quantified. Similar to $\Delta\{\text{pmv-rmv}\}$, the $\Delta\{\text{spmv-rmv}\}$ indicates the range
306 of differences between the correlation matrix entries of the rmv and the sample
307 pseudo mean value chronologies (spmv) of each of the 1000 simulation repetitions.
308 Thus, the spmv are the hypothetical mean value chronologies that result from
309 averaging the phs according to a perfect classification. No further investigations were
310 done on the stability of the poc chronology signals because no further instability
311 results from averaging phs into pocs (for an on-site ratio of 1).

312 **3.4.2 Pref composition**

313 Pref composition was investigated by calculating the percentage of generated series
314 that were classified during each simulation repetition. Of these classified series, the
315 percentage of correctly classified series was determined. To investigate the dynamic
316 establishment of the prefs, the off-site poc/phs contamination was calculated for each
317 run of classifications executed by the PREF-Constructor (section PREF-Constructor
318 Algorithm). That is, the percentage of phs/pocs originating from a pseudo site signal
319 other than the pseudo signal that had been used to initialize the respective pref was
320 calculated in every run.

321 Hypothetically, while the generated prefs may represent mixtures of pseudo site
322 signals they may still carry watershed and/or elevation-specific pseudo signals.
323 Therefore, the phs/pocs were grouped according to contrast groups other than
324 pseudo site provenance. Subsequently, the off-contrast contamination was
325 calculated, i.e. the percentage of phs/pocs originating from a contrast group other

326 than the group that the initial pseudo signal of the respective pref was attributed to.
327 To define these contrast groups, the elevation bands and watersheds/regions of the
328 real site chronologies [15] were used to group the different pseudo site signals:

- 329 • Elevation contrasts (site abbreviations refer to names in [15]; cf. section Real
330 ring-width dataset):
- 331 – Low: real site located < 1000m a.s.l. (sw.pseudo, hw.pseudo, bw.pseudo,
332 ew.pseudo, fri.pseudo, nb.pseudo, sb.pseudo)
- 333 – Medium: real site located 1000-1500m a.s.l. (chw.pseudo, kar.pseudo,
334 how.pseudo, gw.pseudo, gand.pseudo)
- 335 – High: real site located > 1500m a.s.l. (furg.pseudo, rw.pseudo,
336 ww.pseudo)
- 337 • Watershed contrasts:
- 338 – Sihl (pseudo sites: fri.pseudo, furg.pseudo, gw.pseudo, kar.pseudo,
339 sw.pseudo).
- 340 – Linth (gand.pseudo, how.pseudo, hw.pseudo, nb.pseudo, rw.pseudo,
341 ww.pseudo).
- 342 – Obersee (bw.pseudo, chw.pseudo, ew.pseudo, sb.pseudo).

343

4 Results

344

4.1 Stability of the pmv/spmv intercorrelation

345 The range of deviations attributable to variability in the pmv intercorrelation between
346 simulation repetitions, i.e. $\Delta\{\text{pmv-rmv}\}$, was effectively zero, primarily being due to

347 rounding errors ($-5 \cdot 10^{-15}$ to $5 \cdot 10^{-15}$). Moreover, the range of deviations that resulted
 348 from the random sampling of phs across simulation repetitions, i.e. $\Delta\{\text{spmv-rmv}\}$, was
 349 small. The largest positive deviation was 0.039, the largest negative deviation was -
 350 0.083 for any spmv correlation matrix entry.

351 **4.2 Composition of the pseudo reference chronologies**

352 **4.2.1 phs approaches**

353 When setting the t-value threshold to 10 (t10) and 15 (t15), respectively, the
 354 percentage of generated series classified lay below 1.2% with the phs approach
 355 (Table 1). The number of PREF-Constructor runs executed per simulation repetition
 356 was very low (often < 3 runs, Table 1). Only the t5 threshold resulted in prefs that
 357 were long and well replicated enough to allow for further analysis of their
 358 composition.

359 **Table 1: Summary statistics for the different simulation settings (approaches) evaluated.**

approach	% of generated classified			% of classified correct			No. runs executed			No. of series generated		
	Min	Median	Max	Min	Median	Max	Min	Median	Max	Min	Median	Max
poc_t15_1	19.82	34.67	46.33	71.41	96.76	99.93	13	23	39	6322	6417	6501
poc_t15_0.83	4.19	11.47	19.46	61.47	92.70	100.00	3	18	37	6319	6418	6518
poc_t15_0.67	0.82	3.74	8.53	43.53	94.83	100.00	2	9	31	6314	6418	6503
poc_t20_1	0.85	2.85	6.05	75.00	100.00	100.00	2	7	21	6322	6417	6501
poc_t10_1	67.12	78.38	86.23	57.86	75.63	87.15	15	24	47	6322	6417	6501
poc_t10_0.83	56.22	65.56	76.27	53.03	65.52	77.26	13	23	49	6319	6418	6518
poc_t10_0.67	45.88	53.89	60.56	37.35	50.14	61.56	11	17	40	6314	6418	6503
poc_t5_1	83.34	90.54	92.81	30.31	37.38	47.75	7	10	25	6322	6417	6501
phs_t15	0.00	0.00	0.02	0.00	0.00	100.00	0	0	1	6312	6414	6503
phs_t10	0.09	0.44	1.18	77.59	100.00	100.00	1	2	11	6312	6414	6503
phs_t5	43.42	55.54	60.97	33.36	43.28	54.28	11	16	41	6312	6414	6503

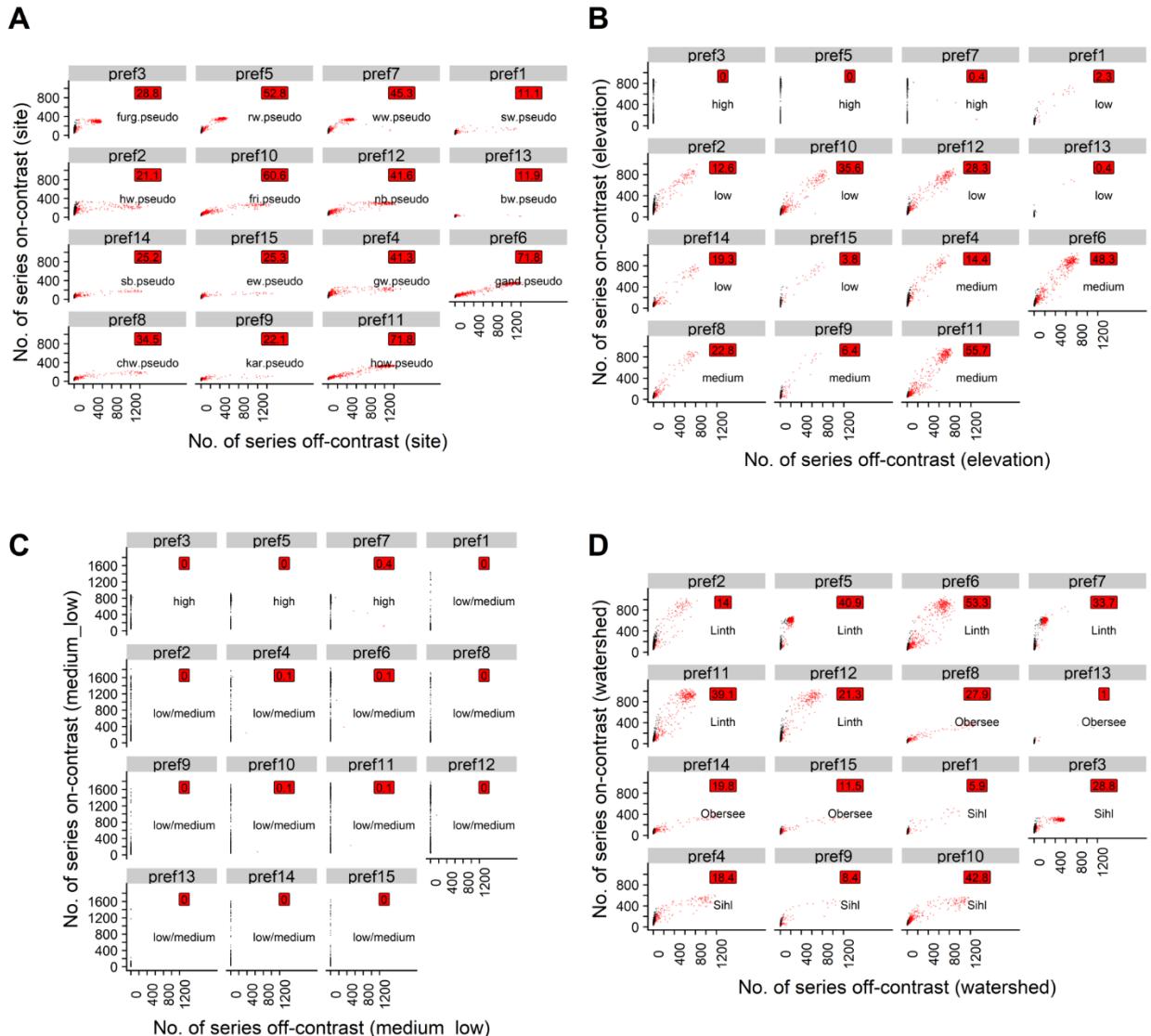
360 However, the t5 prefs were quickly contaminated. By the end of the 2nd PREF-
 361 Constructor run, contamination was <20% for all prefs in 95% of the simulation
 362 repetitions (Table 2). By the end of the final PREF-Constructor run, pref1 and pref13
 363 were the only prefs that were contaminated by less than 20% in more than 80% of
 364 the repetitions (Fig 2A). These were also the shortest and most sparsely replicated
 365 prefs overall (Table 2). Contamination was higher for the longer and better replicated

366 prefs. For example, prefs no. 4, 5, 6, 7, 10 and 11 were contaminated by more than
 367 20% off-site series in >40% of the simulation repetitions (Fig 2A).

368 **Table 2: 95% percentile of contamination, median length, and median mean replication for the**
 369 **phs approach at t5.**

pref1			pref2			pref3			pref4			pref5			
Cont.	Leng.	Repl.													
run1	11.4	218.0	11.1	5.6	231.0	15.1	4.3	231.0	14.9	11.8	231.0	14.3	7.1	230.0	15.6
run2	15.0	245.0	12.6	11.1	288.0	17.7	12.0	286.0	17.5	20.5	283.0	16.6	17.7	298.0	18.6
run3	20.3	257.0	13.6	24.7	342.0	19.7	28.1	330.0	19.7	37.2	317.0	18.7	36.0	352.5	21.3
run4	27.4	264.0	14.3	42.4	368.5	21.1	43.5	357.0	20.8	51.9	338.0	20.0	45.7	395.0	23.1
run5	30.4	265.5	14.6	54.9	388.0	22.0	52.2	370.0	21.4	61.9	347.0	21.1	50.8	421.5	24.3
run6	31.6	267.0	14.8	64.9	402.5	22.5	57.0	374.0	21.7	69.0	353.5	21.4	53.8	432.0	25.6
run7	32.4	267.5	15.0	70.7	411.0	22.7	59.8	377.0	21.9	74.5	355.5	21.6	55.2	432.5	26.6
run8	32.4	268.0	15.0	75.0	417.0	22.8	62.1	378.0	22.0	77.5	358.0	21.8	56.3	437.5	26.8
run9	32.4	268.0	15.0	78.2	422.5	22.9	63.0	378.5	22.0	79.4	358.5	21.9	57.3	437.5	27.0
run10	33.0	268.0	15.1	80.2	427.0	23.0	63.9	379.0	22.0	81.7	359.0	21.9	57.8	437.5	27.1
run11	33.7	268.0	15.1	81.0	430.0	23.1	64.4	379.0	22.0	82.2	359.0	21.9	58.1	437.5	27.1
pref6			pref7			pref8			pref9			pref10			
Cont.	Leng.	Repl.													
run1	16.7	233.0	15.9	7.0	230	15.1	15.6	223.5	12.0	15.0	217	9.9	16.3	229	14.3
run2	35.0	303.0	20.1	18.4	286	18.0	27.1	260.0	13.9	20.8	241	11.3	27.5	282	17.1
run3	59.1	356.5	23.6	36.9	335	20.3	45.0	272.0	15.1	32.8	249	12.1	46.7	313	19.3
run4	68.0	384.5	26.2	46.4	358	21.4	56.2	277.5	15.8	43.3	253	12.5	58.3	327	20.6
run5	72.8	393.5	27.8	52.3	367	22.1	65.8	278.5	16.1	49.2	254	12.8	67.3	329	21.4
run6	75.3	394.0	28.4	54.9	372	22.5	71.2	279.5	16.2	51.7	254	12.9	73.5	332	21.7
run7	76.7	395.0	28.6	56.9	372	22.8	73.4	280.0	16.3	53.1	254	12.9	76.8	335	21.8
run8	77.6	397.5	28.9	58.1	372	22.9	75.4	280.0	16.3	53.7	254	13.0	78.0	335	22.1
run9	78.4	397.5	29.0	59.1	372	23.1	75.0	280.0	16.3	55.2	254	13.1	80.3	335	22.2
run10	78.9	397.5	28.9	59.8	372	23.0	75.1	280.0	16.3	55.2	254	13.1	81.0	335	22.2
run11	79.1	398.5	28.9	60.2	372	23.1	75.5	280.0	16.3	56.1	254	13.1	81.6	335	22.2
pref11			pref12			pref13			pref14			pref15			
Cont.	Leng.	Repl.													
run1	17.0	232.0	15.4	11.1	231.5	15.1	18.2	205.5	6.9	11.1	220.0	12.5	15.2	220.0	11.0
run2	37.2	295.0	19.2	23.7	281.0	18.3	20.6	214.0	7.8	18.4	255.5	14.4	20.0	247.0	12.6
run3	58.7	334.0	22.0	45.3	313.5	20.4	25.0	219.0	8.1	32.6	273.0	15.8	29.5	263.0	13.7
run4	68.8	348.0	23.7	59.9	327.0	21.2	27.0	222.0	8.4	46.9	281.0	16.6	38.5	269.0	14.3
run5	73.5	354.0	24.3	68.6	328.0	21.7	28.6	223.0	8.5	56.3	284.0	17.1	44.5	273.0	14.7
run6	75.8	356.5	24.9	72.9	330.0	21.9	30.3	223.0	8.5	64.8	286.5	17.3	46.1	274.0	15.0
run7	77.6	357.0	25.0	76.4	330.0	22.0	30.6	223.0	8.6	67.1	287.0	17.4	47.0	275.0	15.0
run8	78.7	357.0	25.1	78.2	330.0	22.1	30.8	223.0	8.6	71.2	287.0	17.4	48.1	275.5	15.1
run9	79.4	357.0	25.2	79.6	330.0	22.1	30.8	223.0	8.6	72.0	287.5	17.5	49.2	276.0	15.2
run10	79.8	357.0	25.2	80.4	330.0	22.1	31.1	223.0	8.6	72.7	288.0	17.5	49.5	276.0	15.2
run11	80.0	357.0	25.2	80.8	330.0	22.1	31.5	223.0	8.6	73.3	288.0	17.5	49.5	276.0	15.2

370 Statistics were calculated for the fully replicated PREF-constructor runs only, i.e. runs that were
 371 executed in all 1000 simulation repetitions of the validation process. Full tables including the statistics
 372 for lower replicated runs are provided in the Supporting Information (S3).



373

374 **Fig 2: Scatterplots for the phs approach with t5.** Black points: Uncontaminated simulation
 375 repetitions (<20% off-contrast series). Red points: Contaminated simulation repetitions (>20% off-
 376 contrast series). Red labels: Percentage of simulation repetitions in which the respective pref was
 377 contaminated. Text labels: The contrast group that the respective pref was attributed to. A) Site
 378 contrast; prefs sorted according to elevation band of the real site signals underlying the pseudo site
 379 signals used to initialize the respective prefs. B) Elevation contrast; prefs sorted according to
 380 elevation. C) Elevation contrast rough (medium and low elevations aggregated); prefs sorted
 381 according to elevation. D) Watershed contrast; prefs sorted according to watershed.

382 Regarding contrasts other than the site contrast, prefs initialized with high-elevation
 383 signals only attracted phs sampled from high-elevation pseudo signals. Low- and
 384 medium-elevation signals, however, were often mixed (Fig 2B). The same was true

385 for signals belonging to different watersheds (Fig 2D, S2 Fig 2, S3 Table 1).
 386 Remarkably, in 891 out of 1000 simulation repetitions, 2 full-length prefs were
 387 developed, one of which consisted of high-elevation phs and the other of a mixture of
 388 medium and low-elevation phs. The rest of the prefs were predominantly short (1-333
 389 years) or at least shorter than 667 years (Table 3).

390 **Table 3: Mean number of prefs in the respective length (left) and attractor categories (right).**

approach	length		phs/pocs attracted						
	0	1-333	334-667	668-999	1000	0-100	101-500	501-1000	>1000
poc_t15_1	0	6	3	2	4	8	7	0	0
poc_t15_0.83	0	10	3	1	0	13	2	0	0
poc_t15_0.67	1	13	1	0	0	15	0	0	0
poc_t20_1	3	11	1	0	0	15	0	0	0
poc_t10_1	0	2	5	3	6	1	11	2	0
poc_t10_0.83	0	3	6	2	3	3	9	2	1
poc_t10_0.67	0	7	5	1	2	8	5	1	1
poc_t5_1	0	6	6	1	2	6	6	1	1
phs_t15	15	0	0	0	0	15	0	0	0
phs_t10	7	8	0	0	0	15	0	0	0
phs_t5	0	8	4	0	2	9	4	1	1

391 Attractor category denotes the mean no. of series attracted per pref. For example, in mean (over all
 392 1000 repetitions), there were 8 prefs per repetition that attracted 0-100 pocs with the poc approach at
 393 t15 and osr-1.

394 **4.2.2 poc approaches**

395 • **t15 threshold**

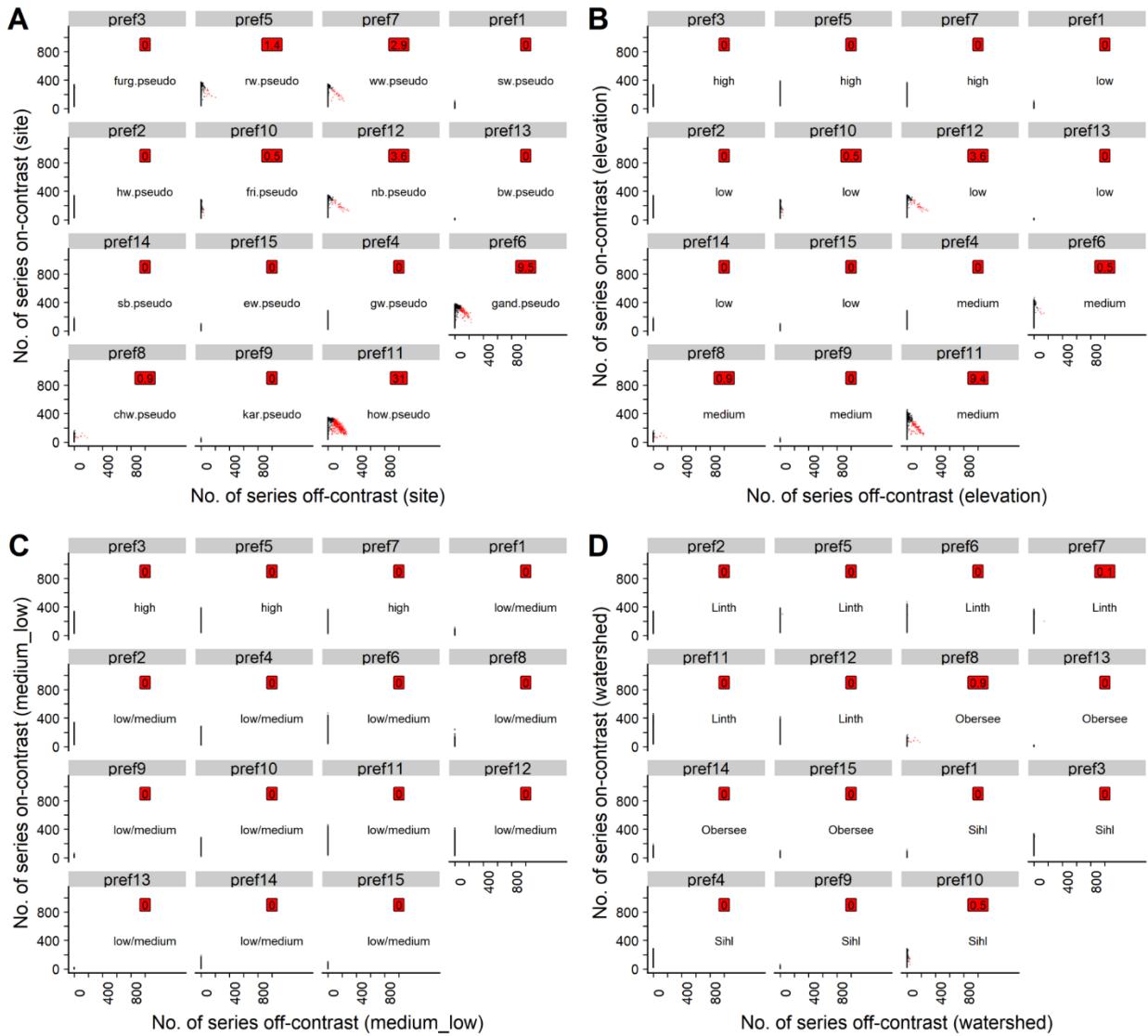
396 When setting the on-site ratio to 1 (osr-1) and the t-value threshold to 15 (t15),
 397 between 19.82% and 46.33% of the generated series were classified by the PREF-
 398 Constructor algorithm. Of these, between 71.41% and 99.93% were classified
 399 correctly (Table 1), depending on the simulation repetition. For those PREF-
 400 Constructor runs that were replicated by all repetitions, the contamination was <20%
 401 until run no. 6 for all prefs generated (Table 4). Moreover, even when calculating the
 402 contamination after the last PREF-Constructor run had been executed, pref6 and

403 pref11 were the only prefs that exhibited a contamination >20% with a notable
404 frequency, i.e. in more than 5% of the repetitions (Fig 3A). For the other prefs, highly
405 contaminated cases (>50%) occurred only rarely (<5% of the repetitions, S3 Table
406 2).

407 **Table 4: 95% percentile of contamination, median length, and median mean replication for the**
 408 **poc approach (t15, osr-1).**

	pref1			pref2			pref3			pref4			pref5		
	Cont.	Leng.	Repl.												
run1	0	205.0	6.6	0	229.0	13.1	0	227.0	13.0	0.0	226.0	11.1	0.0	228.0	14.5
run2	0	221.0	7.5	0	279.5	15.0	0	277.0	15.2	0.0	267.0	12.9	0.0	289.0	17.0
run3	0	228.0	7.9	0	323.5	16.7	0	327.0	16.5	0.0	302.5	14.1	0.0	351.0	18.5
run4	0	230.0	8.2	0	371.0	17.8	0	372.0	17.6	0.0	332.0	15.1	0.0	407.0	19.7
run5	0	232.0	8.4	0	414.0	18.5	0	414.0	18.6	0.0	360.0	15.9	0.6	461.0	20.8
run6	0	232.5	8.5	0	456.5	19.3	0	452.0	19.2	0.0	387.0	16.4	1.0	514.0	21.4
run7	0	232.5	8.6	0	494.0	19.7	0	496.0	19.6	0.0	407.5	16.6	1.4	569.0	21.9
run8	0	232.5	8.6	0	538.0	20.2	0	534.0	20.0	0.0	427.0	16.9	1.6	625.0	22.4
run9	0	232.5	8.7	0	580.5	20.5	0	567.0	20.3	0.0	448.0	17.2	1.8	676.0	22.9
run10	0	232.5	8.7	0	617.0	20.8	0	603.0	20.7	0.5	454.5	17.5	2.3	727.5	23.2
run11	0	232.5	8.7	0	653.5	21.0	0	641.0	20.8	0.5	461.0	17.7	2.4	779.5	23.5
run12	0	232.5	8.7	0	693.0	21.2	0	674.5	21.1	0.5	470.5	17.8	2.5	823.0	23.8
run13	0	232.5	8.7	0	730.0	21.5	0	705.0	21.5	0.5	474.5	17.9	2.8	871.0	24.2
	pref6			pref7			pref8			pref9			pref10		
	Cont.	Leng.	Repl.												
run1	0.0	229.0	14.6	0.0	228.5	13.5	0	211.0	7.8	0	198.0	5.0	0.0	221.0	11.7
run2	0.0	289.0	16.9	0.0	278.0	15.5	0	234.0	8.8	0	210.0	5.7	0.0	263.0	13.4
run3	1.4	348.0	18.7	0.0	331.0	17.1	0	247.0	9.6	0	214.0	6.0	0.0	299.5	14.5
run4	3.0	396.0	20.0	0.0	384.5	18.2	0	255.0	10.2	0	215.0	6.2	0.0	331.0	15.4
run5	5.0	452.5	21.0	0.6	435.0	18.9	0	261.0	10.5	0	215.5	6.3	0.9	357.0	16.2
run6	7.5	508.0	21.6	0.7	476.0	19.8	0	263.5	10.8	0	215.5	6.3	1.0	386.0	16.7
run7	11.0	557.5	22.2	0.8	519.5	20.4	0	264.0	10.9	0	215.5	6.4	1.3	402.0	17.1
run8	14.0	611.0	22.8	1.3	562.5	20.8	0	264.5	11.1	0	215.5	6.4	1.6	424.0	17.4
run9	14.9	659.5	23.2	1.6	598.5	21.3	0	264.5	11.2	0	215.5	6.4	1.7	440.0	17.7
run10	18.5	710.0	23.8	1.8	637.5	21.6	0	264.5	11.2	0	215.5	6.4	1.8	456.5	18.0
run11	22.9	751.0	24.1	2.3	667.0	22.0	0	264.5	11.3	0	215.5	6.4	2.0	468.5	18.2
run12	26.5	792.0	24.4	2.6	698.5	22.1	0	264.5	11.3	0	215.5	6.4	2.2	478.5	18.3
run13	27.9	838.5	24.6	2.8	726.5	22.3	0	264.5	11.3	0	215.5	6.4	2.3	483.5	18.5
	pref11			pref12			pref13			pref14			pref15		
	Cont.	Leng.	Repl.												
run1	0.0	229.0	13.8	0.0	230.5	13.5	0	169.0	2.2	0	210.0	8.3	0	203.0	6.5
run2	0.0	283.0	15.9	0.0	280.0	15.6	0	171.0	2.4	0	234.0	9.4	0	219.0	7.4
run3	1.9	332.0	17.6	0.0	326.5	17.3	0	171.5	2.5	0	254.0	10.2	0	224.0	7.9
run4	5.0	376.5	19.0	1.0	374.5	18.2	0	171.5	2.5	0	269.5	10.8	0	228.0	8.2
run5	10.1	420.0	19.7	1.6	418.0	19.2	0	171.5	2.5	0	275.5	11.2	0	229.5	8.4
run6	17.9	460.5	20.6	2.5	457.5	19.7	0	171.5	2.5	0	278.0	11.5	0	230.0	8.5
run7	24.6	496.0	21.1	2.8	487.5	20.4	0	171.5	2.5	0	278.0	11.6	0	230.0	8.6
run8	31.6	524.0	21.5	3.2	530.0	20.6	0	171.5	2.5	0	278.0	11.9	0	230.0	8.6
run9	38.0	553.5	21.8	3.4	569.0	21.0	0	171.5	2.5	0	278.0	12.0	0	230.0	8.6
run10	42.1	579.0	22.2	4.3	603.0	21.4	0	171.5	2.5	0	278.0	12.0	0	230.0	8.7
run11	45.2	597.5	22.4	5.6	634.0	21.6	0	171.5	2.5	0	278.5	12.1	0	230.0	8.7
run12	47.2	611.5	22.6	6.4	663.5	21.9	0	171.5	2.5	0	278.5	12.1	0	230.0	8.7
run13	49.5	625.0	22.8	7.6	690.5	22.1	0	171.5	2.5	0	278.5	12.1	0	230.0	8.7

409 Statistics were calculated for fully replicated PREF-constructor runs only, i.e. runs that were executed
 410 in all 1000 simulation repetitions of the validation process. Full tables including statistics for lower
 411 replicated runs are provided in the Supporting Information (S3).



412

413 **Fig 3: Scatterplots for the poc approach with t15 and osr-1.** Black points: Uncontaminated
 414 simulation repetitions (<20% off-contrast series). Red points: Contaminated simulation repetitions
 415 (>20% off-contrast series). Red labels: Percentage of simulation repetitions in which respective pref
 416 was contaminated. For text labels and explanation of the subplots, cf. caption of Fig 2.

417 On average, per repetition 4 prefs reached the full length of 1000 years. The majority
 418 of the prefs were >333 years and often >500 years long. However, in each repetition,
 419 on average 6 of the prefs reached <334 years (Table 3, S3 Table 2). The ranking of
 420 shortest to longest pref tended to differ considerably by repetition (e.g. pref1 in
 421 certain repetitions was among the longest and in other repetitions among the shortest
 422 prefs). However, over all simulation repetitions, certain prefs developed

423 systematically better than others (Table 4, S2 Fig 3, S3 Table 2). Among the
424 frequently longer prefs were no. 2, 3, 5, 6 and 7, with a median length >700 years
425 (calculated at the last PREF-constructor run replicated by all 1000 simulation
426 repetitions, Table 4). The shorter prefs were no. 1, 4, 8, 9, 10, 13, 14 and 15, as
427 reflected by a generally lower median length (<500 years; calculated at the last fully
428 replicated PREF-constructor run, Table 4). These prefs also had a lower replication,
429 i.e. a median mean replication <20 series by the end of the last (13th) fully replicated
430 PREF-Constructor run (Table 4; for full tables of minimum, median and maximum
431 lengths and mean replication, respectively, cf. S3 Table 2).

432 Lowering the on-site ratio to 0.83 (osr-0.83) and 0.67 (osr-0.67), respectively,
433 resulted in a drop of classification rates (Table 1). Consequently, fewer PREF-
434 Constructor runs were executed per repetition. Especially with osr-0.67, prefs
435 generally remained short and sparsely replicated (S2 Figs 4-7, S3 Table 3-4). In
436 extreme cases, e.g. pref13 initialized with bw.pseudo, prefs remained quasi in their
437 initial state (S2 Figs 6-7, S3 Table 4, S4 Text).

438 • **t20, t10 and t5 threshold**

439 Short and sparsely replicated prefs resulted also when increasing the t-value
440 threshold to t20 at osr-1, i.e. the maximum on-site ratio (S2 Figs 8-9, S3 Table 5, S4
441 Text). As for t15 at osr-0.67, prefs often remained in their initial state. Thus, no
442 simulations featuring alternative on-site ratios were executed for t20.

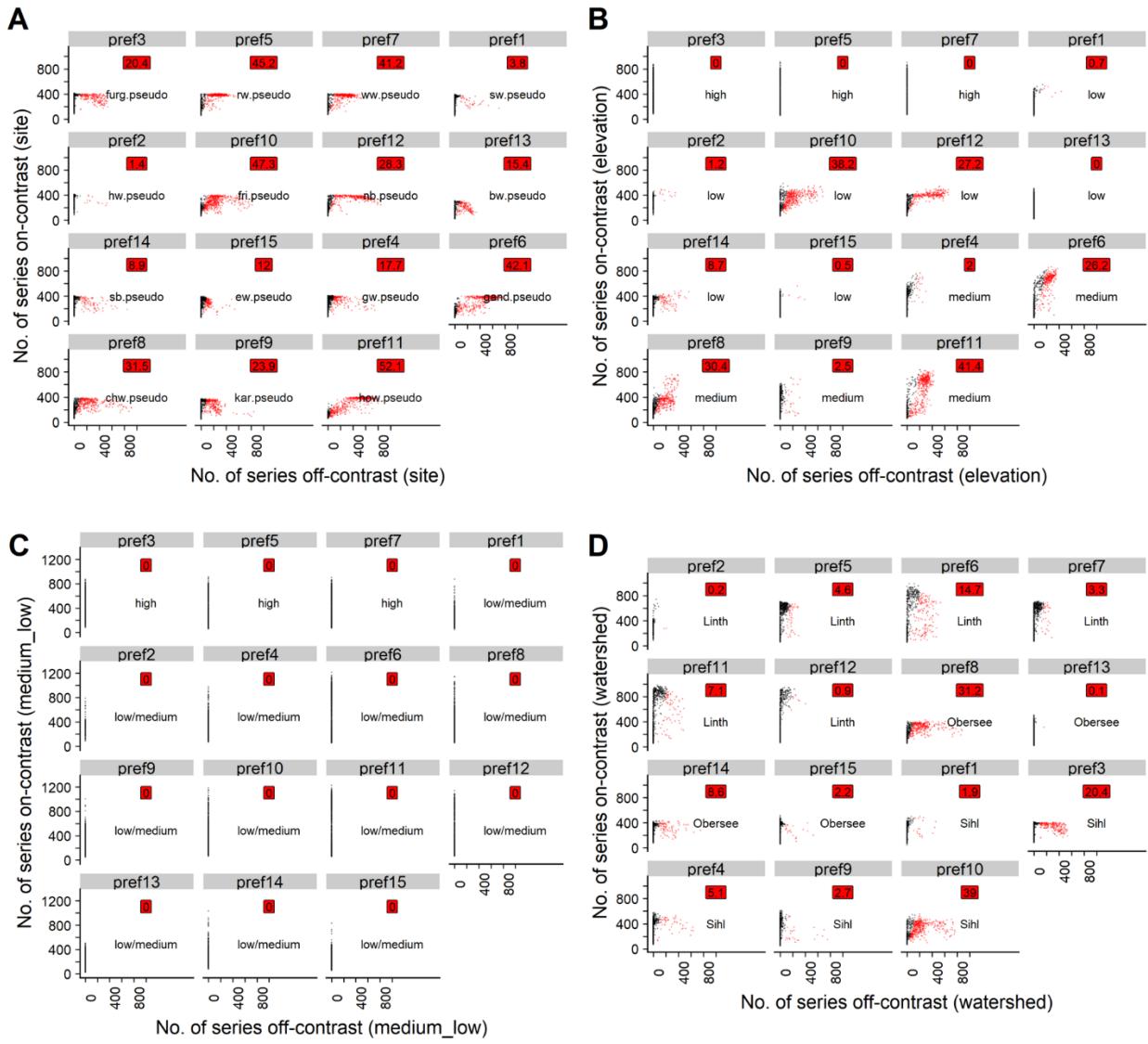
443 For t10 at osr-1, the contamination risk varied considerably between prefs. After the
444 first three PREF-Constructor runs, in 95% of the repetitions most prefs, i.e. no. 1, 2,
445 3, 4, 5, 7, 8, 9, 10, 13, 14 and 15, exhibited a low contamination (<20%, Table 5).
446 However, in the bulk of repetitions, the majority of prefs were contaminated
447 progressively with every subsequent PREF-Constructor run (Table 5). Nevertheless,

448 usually long (median 453 to 1000 years) and well replicated (median 20 to 28.7)
449 prefs were established with t10 at osr-1 (medians calculated at PREF-constructor run
450 15, Tables 3 and 5, Fig 4, S3 Table 6, S4 Text, S5 Fig 10). Decreasing the on-site
451 ratio for t10 did not affect the length and replication of the prefs as heavily as for t15
452 (Tables 1 and 3). In summary however, the same observations were made, i.e. the
453 number of correctly classified series decreased the lower the on-site ratio (Table 1,
454 Figs 5-6, S3 Tables 7-8, S5 Figs 11-12). Consequently, and especially with osr-0.67,
455 the length and replication of prefs decreased in most repetitions (S4 Text).

456 **Table 5: 95% percentile of contamination, median length, and median mean replication for the**
 457 **poc approach (t10, osr-1).**

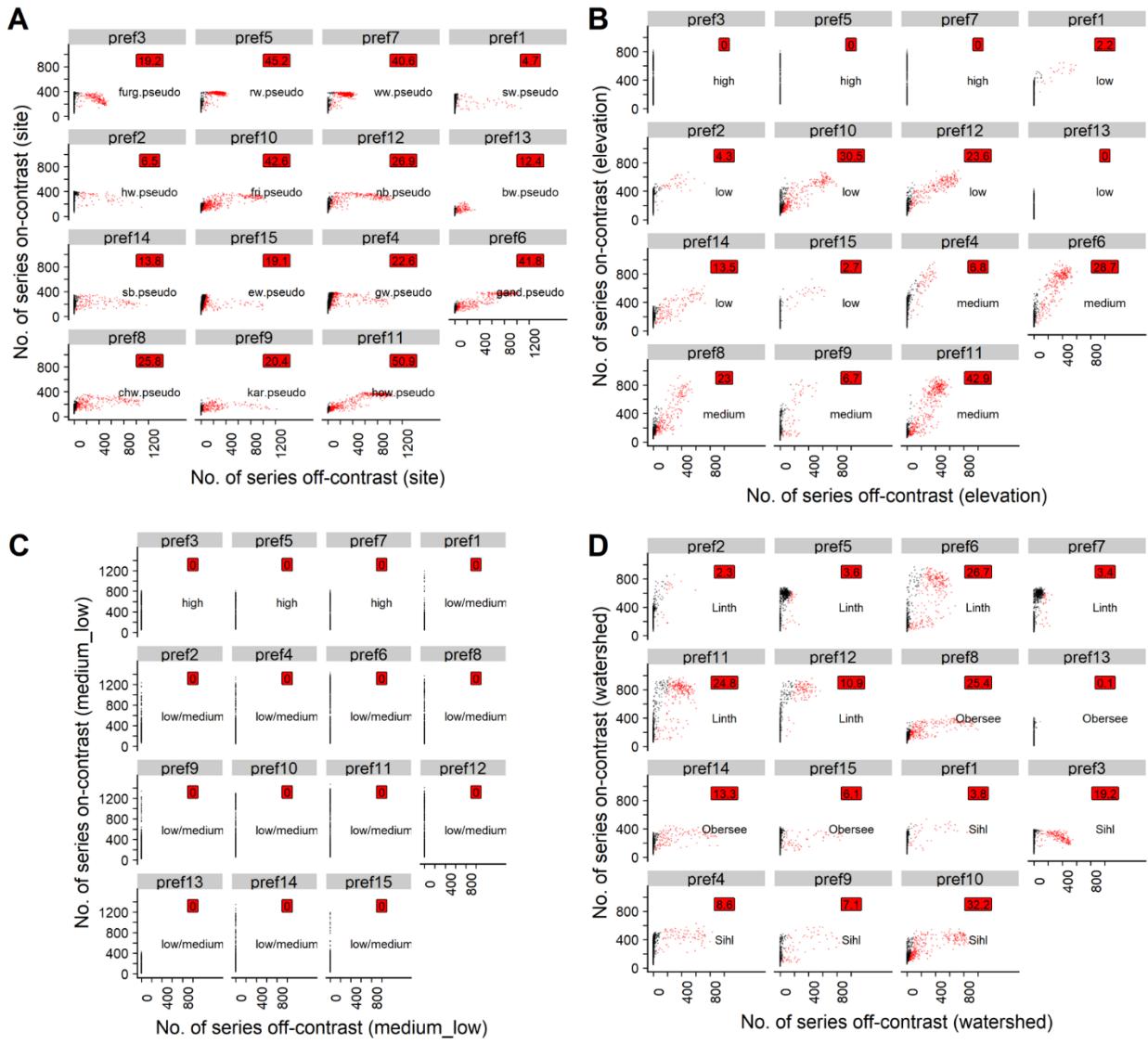
	pref1			pref2			pref3			pref4			pref5		
	Cont.	Leng.	Repl.												
run1	0.0	233.0	15.3	0.0	235	16.1	0.0	234.0	16.1	0.0	235.0	16.0	0.0	232.0	16.2
run2	0.0	299.0	17.9	0.0	310	19.7	1.2	315.0	19.8	1.6	312.0	19.6	5.0	318.0	19.8
run3	1.5	369.0	19.8	0.5	393	21.8	3.5	400.0	21.7	4.7	390.0	21.7	16.8	403.0	22.1
run4	2.2	424.0	21.3	0.6	467	23.1	8.8	481.0	23.3	9.4	464.0	23.1	28.6	481.5	23.9
run5	3.6	473.0	22.4	0.8	537	24.1	18.6	546.0	24.4	14.7	528.0	24.1	36.3	553.0	24.8
run6	5.4	519.0	23.0	1.0	609	24.7	30.5	614.5	24.9	19.5	586.5	25.1	41.9	611.5	25.8
run7	7.5	565.0	23.6	1.3	680	25.3	40.3	671.0	25.5	27.0	630.5	25.6	44.1	669.5	26.2
run8	8.3	598.5	24.0	1.3	738	25.7	46.9	725.0	25.9	31.3	682.0	26.2	46.3	713.0	26.7
run9	9.1	637.0	24.3	1.3	799	26.1	51.6	769.5	26.2	34.5	718.5	26.5	47.6	745.5	27.2
run10	9.3	665.5	24.6	1.3	859	26.3	54.3	803.5	26.5	36.5	756.5	26.8	48.1	760.0	27.8
run11	9.1	698.0	24.7	1.3	913	26.7	54.8	828.0	26.6	37.2	777.0	27.2	48.8	777.5	28.3
run12	10.8	726.5	24.9	1.3	949	27.0	55.9	838.5	26.8	38.2	816.0	27.4	48.9	782.5	28.5
run13	10.8	756.5	24.9	1.3	1000	27.5	56.4	850.0	26.9	38.9	843.0	27.6	49.0	783.5	28.6
run14	10.8	785.5	25.0	1.3	1000	28.2	56.3	866.5	27.0	39.6	872.0	27.8	49.0	784.0	28.7
run15	10.8	818.0	25.1	1.3	1000	28.3	56.3	876.5	27.0	40.3	888.5	28.0	49.1	784.0	28.7
	pref6			pref7			pref8			pref9			pref10		
	Cont.	Leng.	Repl.												
run1	0.0	232.0	16.3	0.0	232.0	16.2	0.0	233.0	15.6	0.0	234.0	14.7	0.0	231.0	16.3
run2	12.2	312.0	20.3	4.9	312.0	19.8	3.7	301.5	18.8	1.4	294.0	17.4	4.2	312.0	19.6
run3	29.8	383.0	22.7	18.3	395.5	22.1	14.3	365.5	20.9	3.8	348.5	19.3	14.2	380.0	22.1
run4	41.3	440.0	24.0	30.8	463.0	23.7	26.7	414.5	22.5	11.3	400.0	20.6	27.4	444.0	23.6
run5	50.7	472.5	25.0	37.8	528.0	24.8	36.6	455.0	23.6	22.0	439.0	21.7	37.5	487.0	24.7
run6	55.7	491.0	25.5	43.8	584.5	25.4	43.6	482.5	24.2	30.6	475.0	22.3	43.5	522.0	25.6
run7	59.6	497.0	25.8	46.6	621.0	25.9	48.4	504.0	24.6	39.2	506.0	22.9	50.1	543.0	26.2
run8	61.7	501.0	26.0	49.0	650.0	26.3	51.6	517.0	24.9	43.9	529.0	23.2	54.3	563.5	26.4
run9	62.6	501.0	26.0	49.4	665.0	26.4	54.9	526.5	25.1	47.4	551.0	23.5	58.2	581.5	26.7
run10	63.2	503.5	26.2	50.2	668.0	26.7	57.6	539.0	25.2	48.6	575.0	23.6	60.2	598.0	26.9
run11	63.4	503.5	26.2	50.2	668.5	26.8	58.0	541.5	25.2	51.0	589.0	23.8	61.3	613.5	27.1
run12	63.5	503.5	26.2	50.3	668.5	26.9	58.5	546.5	25.2	53.4	606.5	23.8	61.7	620.5	27.3
run13	63.6	503.5	26.3	50.3	668.5	27.0	58.9	550.0	25.2	53.7	616.0	23.9	62.0	630.0	27.2
run14	63.7	503.5	26.3	50.3	668.5	27.0	58.9	552.5	25.2	54.5	624.5	24.0	62.1	631.0	27.4
run15	63.6	503.5	26.3	50.4	668.5	27.0	59.1	552.5	25.3	54.6	637.0	24.0	62.1	632.0	27.5
	pref11			pref12			pref13			pref14			pref15		
	Cont.	Leng.	Repl.												
run1	0.0	234.0	16.2	0.0	235.0	16.1	0.0	225.0	12.5	0.0	233.0	15.8	0.0	230.0	15.4
run2	16.6	314.0	20.3	8.0	312.0	20.0	0.0	267.5	14.3	0.0	301.0	18.8	1.0	293.0	18.1
run3	38.3	384.0	22.9	21.1	390.0	22.3	1.8	303.0	15.8	1.6	370.5	20.9	2.0	357.0	19.8
run4	47.7	446.0	24.5	34.8	453.0	23.7	4.0	333.0	16.8	2.7	431.0	22.2	4.3	417.0	21.4
run5	53.6	487.5	25.9	46.2	492.5	24.5	7.7	359.5	17.7	4.7	488.0	23.2	6.3	473.5	22.3
run6	57.8	508.5	27.2	52.2	526.0	25.1	15.5	381.0	18.2	8.8	529.0	24.1	9.5	526.5	23.2
run7	59.8	516.0	27.9	56.4	543.0	25.4	20.3	400.5	18.6	13.4	571.5	24.5	12.1	572.0	23.9
run8	61.6	521.0	28.3	60.2	550.5	25.7	25.3	414.5	18.8	19.5	605.5	24.9	13.4	610.0	24.4
run9	62.7	521.5	28.6	61.6	556.0	25.8	27.8	425.5	19.1	23.4	640.0	25.0	16.0	649.0	24.8
run10	63.3	521.5	28.7	61.9	558.0	25.8	30.3	433.0	19.3	29.6	671.0	25.3	17.8	683.0	25.2
run11	63.6	522.5	28.7	61.9	561.0	25.9	34.1	439.5	19.5	33.6	701.0	25.3	20.2	717.5	25.5
run12	63.7	522.5	28.7	61.9	564.0	25.9	36.4	443.0	19.7	35.7	728.5	25.5	22.7	756.5	25.7
run13	63.8	522.5	28.7	61.9	564.0	25.9	38.3	448.0	19.8	37.1	754.5	25.4	24.2	800.0	25.8
run14	64.0	522.5	28.7	61.9	564.0	25.9	40.8	449.0	19.9	37.4	780.0	25.5	24.9	824.0	26.1
run15	64.1	522.5	28.7	61.9	564.0	25.9	42.5	453.0	20.0	38.4	809.0	25.5	26.5	865.5	26.4

458 Statistics were calculated for fully replicated PREF-constructor runs only, i.e. runs that were executed
 459 in all 1000 simulation repetitions of the validation process. Full tables including statistics for lower
 460 replicated runs are provided in the Supporting Information (S3).



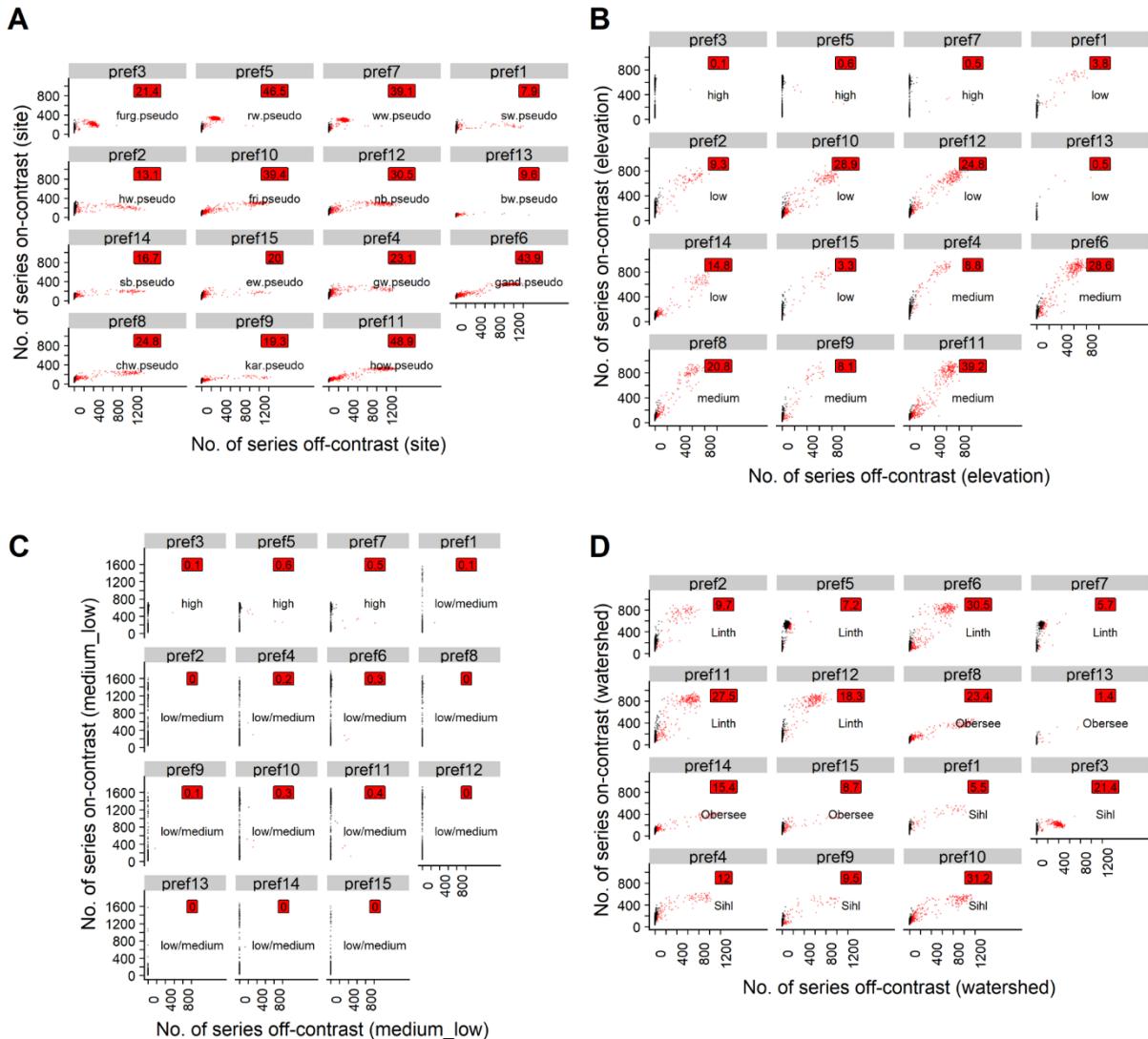
461

462 **Fig 4: Scatterplots for the poc approach with t10 and osr-1.** Black points: Uncontaminated
 463 simulation repetitions (<20% off-contrast series). Red points: Contaminated simulation repetitions
 464 (>20% off-contrast series). Red labels: Percentage of simulation repetitions in which the respective
 465 pref was contaminated. For text labels and explanation of the subplots, cf. caption of Fig 2.



466

467 **Fig 5: Scatterplots for the poc approach with t10 and osr-0.83.** Black points: Uncontaminated
 468 simulation repetitions (<20% off-contrast series). Red points: Contaminated simulation repetitions
 469 (>20% off-contrast series). Red labels: Percentage of simulation repetitions in which the respective
 470 pref was contaminated. For text labels and explanation of the subplots, cf. caption of Fig 2.



471

472 **Fig 6: Scatterplots for the poc approach with t10 and osr-0.67.** Black points: Uncontaminated
 473 simulation repetitions (<20% off-contrast series). Red points: Contaminated simulation repetitions (>
 474 20% off-contrast series). Red labels: Percentage of simulation repetitions in which the respective pref
 475 was contaminated. For text labels and explanation of the subplots, cf. caption of Fig 2.

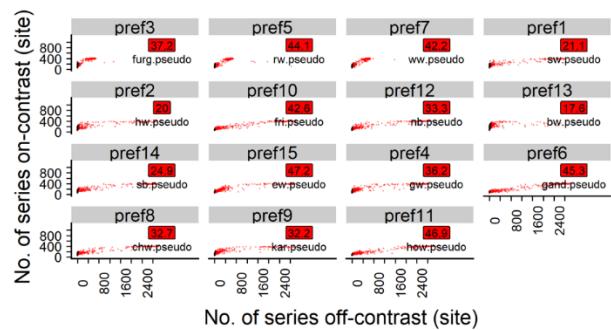
476 A very high contamination risk was observed when lowering the threshold to t5, even
 477 while keeping the on-site ratio at its maximum, i.e. osr-1 (Tables 1, 3 and 6, Fig 7, S3
 478 Table 9, S4 Text, S5 Fig 13). Thus, for the poc approach at t5, the simulations
 479 featuring lower on-site ratios were omitted.

480 **Table 6: 95% percentile of contamination, median length, and median mean replication for the**
 481 **poc approach (t5, osr-1).**

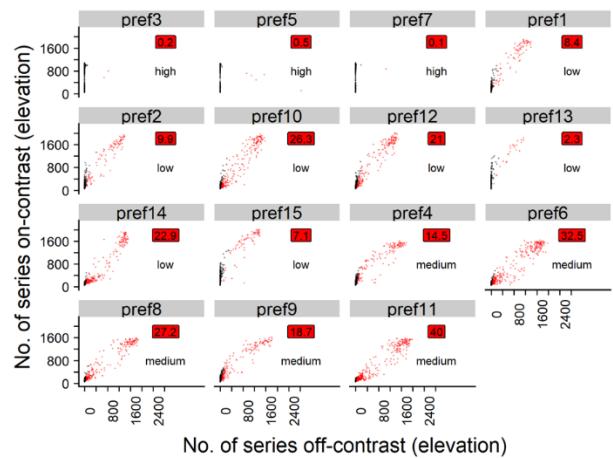
pref1			pref2			pref3			pref4			pref5			
Cont.	Leng.	Repl.													
run1	0.0	235	16.1	0.0	235.0	16.2	0.0	234.0	16.1	0.0	235	16.0	0.0	232.0	16.2
run2	35.8	317	20.9	32.2	313.0	20.6	35.0	323.5	20.7	45.8	321	21.2	38.8	329.0	20.9
run3	60.0	356	22.2	59.7	369.5	22.0	49.7	397.0	22.2	69.9	356	22.4	52.1	398.5	22.6
run4	72.5	359	22.3	74.0	373.5	22.3	55.8	435.0	22.8	79.0	357	22.8	57.2	419.5	23.6
run5	78.7	359	22.3	79.8	373.5	22.4	58.5	443.5	22.9	83.3	357	22.8	59.5	421.0	23.8
run6	83.2	359	22.3	83.9	373.5	22.4	60.1	443.5	23.0	85.4	357	22.8	60.9	421.0	23.8
run7	84.1	359	22.4	84.7	373.5	22.4	61.1	443.5	23.0	86.2	357	22.8	61.8	421.0	23.8
pref6			pref7			pref8			pref9			pref10			
Cont.	Leng.	Repl.													
run1	0.0	232	16.3	0.0	232.0	16.2	0.0	233.5	16.0	0.0	236.0	16.0	0.0	231	16.3
run2	63.4	322	21.9	37.6	324.0	20.9	50.0	315.0	21.6	40.7	316.0	21.4	57.3	319	21.6
run3	77.3	343	23.2	51.1	389.0	22.8	74.1	340.0	22.5	62.4	346.5	22.7	75.1	342	22.8
run4	82.9	343	23.4	56.7	401.0	23.4	81.2	340.0	22.5	73.9	349.0	23.0	82.0	343	23.1
run5	85.6	343	23.4	58.9	403.5	23.6	84.4	340.0	22.6	81.1	350.0	23.1	84.6	343	23.1
run6	86.9	343	23.5	60.3	403.5	23.6	86.1	340.0	22.6	84.4	350.0	23.1	86.3	343	23.1
run7	87.3	344	23.5	61.4	403.5	23.6	86.7	340.0	22.6	84.6	350.0	23.1	86.9	343	23.1
pref11			pref12			pref13			pref14			pref15			
Cont.	Leng.	Repl.													
run1	0.0	234.0	16.2	0.0	235.0	16.1	0.0	235.0	16.0	0.0	233.0	16.1	0.0	231.5	16.3
run2	60.6	323.5	22.1	50.4	318.5	21.3	22.5	311.0	20.6	46.5	312.0	20.9	40.9	314.0	21.4
run3	77.1	343.5	23.6	72.3	358.0	22.5	38.5	366.5	22.5	70.8	339.5	22.1	63.0	354.0	23.1
run4	82.8	344.0	23.9	80.9	358.0	22.6	47.9	381.0	23.1	80.3	340.0	22.3	71.8	363.5	23.8
run5	85.6	344.0	23.9	84.3	358.0	22.7	52.4	388.0	23.3	83.8	340.0	22.3	79.6	369.5	23.9
run6	87.0	344.0	23.9	86.0	358.0	22.7	53.4	390.0	23.3	86.0	340.0	22.3	82.3	370.0	24.0
run7	87.3	344.0	23.9	86.6	358.0	22.7	52.4	391.0	23.4	86.3	340.0	22.3	82.7	370.0	24.1

482 *Statistics were calculated for fully replicated PREF-constructor runs only, i.e. runs that were executed*
 483 *in all 1000 simulation repetitions of the validation process. Full tables including statistics for lower*
 484 *replicated runs are provided in the Supporting Information (S3).*

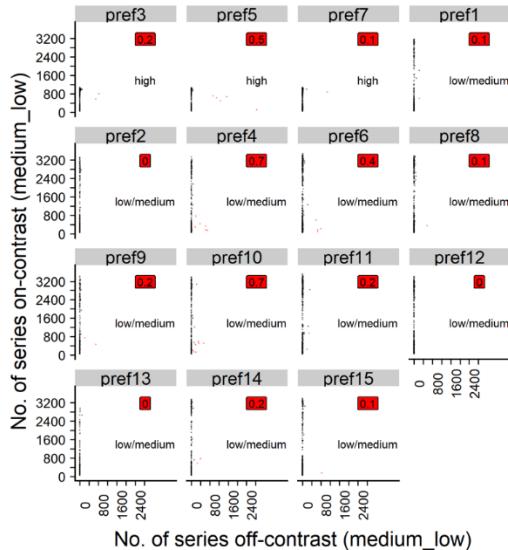
A



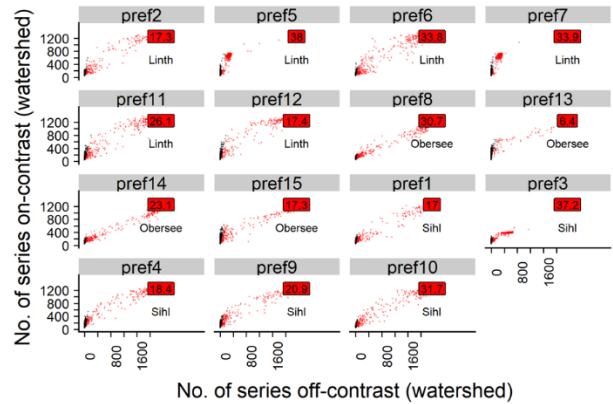
B



C



D



485

Fig 7: Scatterplots for the poc approach with t5 and osr-1. Black points: Uncontaminated simulation repetitions (<20% off-contrast series). Red points: Contaminated simulation repetitions (>20% off-contrast series). Red labels: Percentage of simulation repetitions in which respective pref was contaminated. For text labels and explanation of the subplots, cf. caption of Fig 2.

For all t-value thresholds and on-site ratios investigated, grouping the pseudo signals according to the elevation bands of the underlying real sites revealed that prefs initialized with pseudo signals from high elevation sites did generally not attract poc series that were sampled from medium- or low-elevation pseudo signals (Figs 4C, 5C, 6C and 7C). However, prefs frequently attracted pocs from watersheds other

495 than the pseudo site signal used to initialize the respective pref (Figs 4D, 5D, 6D and
496 7D).

497 **5 Discussion**

498 **5.1 A simulation model for investigating the contamination 499 risk of reference chronologies**

500 The aim of this study was to investigate site signal preservation based on a
501 simulation approach. The small range of deviances between the correlation matrix
502 entries of the real site chronologies and the sample pseudo mean value chronologies
503 ($\Delta\{\text{spmv-rmv}\}$) indicated that the simulation model produced pseudo chronologies
504 with a suitably stable intercorrelation structure of the underlying pseudo site signals.
505 Moreover, a low contamination risk was attested to the approach using unmixed (on-
506 site ration of 1, i.e. osr-1) pseudo object chronologies (poc/pocs) and the t-value
507 threshold set to $t \geq 15$ (t15). This demonstrates that the PREF-Constructor algorithm
508 is capable of establishing prefs that preserve the pseudo site signal. This result is
509 most relevant for dendro-provenancing as it provides evidence for the adequacy of
510 the theoretical concept of reference chronologies. Under ideal circumstances, long
511 reference chronologies that incorporate historical series have the potential to
512 represent local tree growth. To date, no other study seems to have tried to
513 substantiate this key assumption of reference chronologies [5–14,18–21].

514 **5.2 Implications and benefits for dendro-provenancing**

515 Only the poc approach with t15 and osr-1 provided ideal conditions for developing
516 uncontaminated prefs. Lowering the osr for the poc approach at t15 diluted their

517 pseudo site signal and resulted in a quick decrease in pref length and replication.
518 This often prohibited the development of longer prefs for weaker pseudo site signals
519 (e.g. bw.pseudo that was used to initialize pref13, cf. [15]).

520 In the context of sampling real timbers from historical objects, an osr of 1 seems too
521 optimistic [33]. Realistically, imported timbers occasionally will enter object
522 chronologies [39]. Unfortunately, the simulation was not robust when off-site series
523 were incorporated in pocs. Hence, researchers are best advised to scrutinize the
524 assumption of local timber supply for each object investigated when attempting to
525 establish object chronologies for dendro-provenancing. In addition, establishing
526 reference chronologies using single ring-width series bears a high contamination risk,
527 as indicated by the quickly contaminated prefs established with the phs approach at
528 t5. In the median, three quarters of the phs were either unclassified or wrongly
529 classified (Table 1). Moreover, the pseudo site signal was too weak in the phs, thus
530 the threshold could not be raised above t5. Hence, at least for small-scale studies
531 like [15], establishment of reference chronologies and provenancing based on single
532 ring-width series is likely to be highly problematic.

533 A t-value threshold of 15 is very restrictive. In the literature, t15 is an uncommon
534 threshold: often, researchers applied more or less rigorously fixed thresholds
535 between t9 and t11 [7,8,40–43]. Some studies even considered t-values <9 to narrow
536 down the area of provenance [44,45]. Strictly speaking, the t-value thresholds are not
537 directly comparable between different studies because the t-value is influenced by
538 preprocessing and the overlap over which the correlation coefficient is calculated
539 [46,47]. However, the most commonly used t-value calculation methods involve
540 transformations that enhance the high-frequency signal [17,18]. Thus, the t-value for
541 a correlation coefficient still provides a good approximation of the similarity between

542 ring-width series that was deemed sufficient for dendro-provenancing in the
543 respective studies cited above (i.e., [7,8,40–45]). The simulation approach that was
544 developed in the present study shows that, even with t15 and osr-1, provenancing
545 was successful for only one third of the pocs in median (i.e. 34.67% classified, of
546 those 96.76% correct; Table 1). For the present dataset, lower thresholds evinced a
547 higher risk of establishing contaminated prefs. For example, although the t10
548 threshold exhibited a low contamination risk for establishing prefs in the majority of
549 the repetitions of this simulation study, there were several repetitions in which prefs
550 that were uncontaminated in most other repetitions were severely contaminated.
551 Particularly, variability was extremely high among the 5% most contaminated
552 simulation repetitions. Thus, no t-threshold can be considered “secure”. Even for t15
553 with osr-1, contaminated prefs resulted in rare cases.

554 A simulation approach offers the means for assessing the level of statistical similarity
555 necessary for lowering the risk of establishing contaminated prefs within a given
556 dataset. The t-thresholds determined here are neither universally valid nor directly
557 transferable to other study regions. Because computational and hardware resources
558 were limited, thresholds had to be set for t-values and overlap. Preferably, the
559 calculation would have been continuous. In any real dendro-provenancing study, the
560 dataset displays other clusters of between ring-width series similarities [48–51]. In
561 ideal circumstances, the between-sites signal differences in an area are pronounced
562 enough to allow for site-specific similarity clusters. The range of t-value thresholds
563 that capture the respective level of similarity necessary to establish site-specific,
564 elevation-specific or regional reference chronologies, respectively, should be
565 investigated independently for each study region. This is only possible via simulation,
566 but it seems to have been largely disregarded up to now [5–14].

567 Moreover, thresholds or t-value ranges that were determined based on simulation
568 require re-evaluation when expanding the dataset of a study region. On the one
569 hand, the pseudo signals studied here appear disparate when applying the t15
570 threshold. For neighboring sites, however, narrowing the sampling grid by adding
571 new site chronologies may gradually dissolve the between-site signal differences. On
572 the other hand, certain signal differences may turn out to be more robust and may
573 persist even when new data are added. For example, pronounced elevation specific
574 growth signals have been reported by several studies [52–57]. In the simulation
575 presented here, high-elevation and medium- to low-elevation pseudo site signals
576 proved robust and were separated even with the phs approach at t5 or the poc
577 approach at t10 and osr-0.67.

578 **6 Conclusions**

579 Simulation is paramount for future dendro-provenancing studies. In the dataset
580 studied here, uncontaminated prefs were established only with very strict settings for
581 the t-value and osr thresholds. In any study region, the contamination risk of local
582 reference chronologies is unknown *a priori* as the site provenance of historical
583 timbers remains ambiguous. Thus, this risk can be objectively assessed only through
584 simulation.

585 The approach presented here can be extended in several ways: For example, the
586 input covariance matrix of the real sites could be replaced by any covariance matrix.
587 Also, the thresholds for t-values, on-site ratio and overlap may be replaced by
588 continuous inputs. However, such developments require high-performance
589 programming techniques and/or hardware. These were beyond the resources
590 available to the author and also beyond the scope of this paper.

591 Another valuable path for future research would focus on validating or revising the
592 simulation model presented here. For all of the 1000 simulated years, the same
593 intercorrelation structure between the pseudo site signals was assumed. In reality,
594 some temporal variability in the intercorrelation between sites is likely to occur, which
595 could be reflected in the model but was ignored here for the sake of simplicity.

596 In addition to elaborating the simulation model, it could be extended to incorporate
597 multivariate tree-ring time series of proxies such as wood-density. Multivariate
598 approaches were shown to increase cross-dating success [58,59]. Thus, such an
599 approach is likely to enhance the spatial resolution of dendro-provenancing and
600 lower the risk of local reference chronology contamination.

601 Without an adequate evaluation of the contamination risk of reference chronologies,
602 the basis of dendro-provenancing remains inscrutable. The approach presented here
603 is a first and not necessarily the most practicable solution. The depicted avenues of
604 future research hopefully spur methodological innovations in the field that will lead to
605 a more elaborated simulation tool for practical application in dendro-provenancing.

606 **7 Acknowledgements**

607 The author thanks H. Bugmann and N. Bleicher for comments and suggestions on
608 and the discussions of the manuscript; and H. Bugmann and Ph. Della Casa for
609 supervising the author's PhD thesis.

610 **8 References**

- 611 1. Speer JH. Fundamentals of Tree-Ring Research. Tucson: University of Arizona
612 Press; 2010.

- 613 2. Bleicher N. Pastoralism, Landscape and Woodland Use in the Late Neolithic in

614 Southern Germany. In: Heiss AG, Retamero F, Schjellerup I, Davies A,

615 editors. Agricultural and Pastoral Landscapes in Pre-Industrial Society: Oxbow

616 Books; 2014. pp. 239-250.

617 3. Billamboz A. Regional patterns of settlement and woodland developments:

618 Dendroarchaeology in the Neolithic pile-dwellings on Lake Constance

619 (Germany). Holocene. 2014;24: 1278-1287. doi:10.1177/0959683614540956.

620 4. Bleicher N, Harb C. Settlement and social organisation in the late fourth

621 millennium BC in Central Europe: the waterlogged site of Zurich-Parkhaus

622 Opéra. Antiquity. 2018;92: 1210-1230. doi: 10.15184/aqy.2018.109.

623 5. Drake BL. Source & Sourceability: Towards a probabilistic framework for

624 dendroprovenance based on hypothesis testing and Bayesian inference.

625 Dendrochronologia. 2018;47: 38-47. doi: 10.1016/j.dendro.2017.12.004.

626 6. Domínguez-Delmás M, van Daalen S, Alejano-Monge R, Wazny T. Timber

627 resources, transport and woodworking techniques in post-medieval Andalusia

628 (Spain): Insights from dendroarchaeological research on historic roof

629 structures. J Archaeol Sci. 2018;95: 64-75. doi: 10.1016/j.jas.2018.05.002.

630 7. Bernabei M, Bontadi J. Determining the resonance wood provenance of stringed

631 instruments from the Cherubini Conservatory Collection in Florence, Italy. J

632 Cult Herit. 2011;12: 196-204. doi: 10.1016/j.culher.2010.12.001.

- 633 8. Eckstein D, Wrobel S. Dendrochronological proof of origin of historic timber -
634 retrospect and perspectives. In: Haneca K, Verheyden A, Beekmann H,
635 Gärtner H, Helle G, Schleser G, editors. TRACE - Tree Rings in Archaeology,
636 Climatology and Ecology, Vol.5: Proceedings of the Dendrosymposium 2006,
637 April 20th - 22nd 2006, Tervuren, Belgium; 2007. pp. 8-20.
- 638 9. Eissing T, Dittmar C. Timber transport and dendroprovenancing in Thuringia and
639 Bavaria. In: Fraiture P, editor. Tree Rings, Art and Archaeology; 2011. pp.
640 137-149.
- 641 10. Domínguez-Delmás M, Driessen M, García-González I, van Helmond N, Visser
642 R, Jansma E. Long-distance oak supply in mid-2nd century AD revealed: the
643 case of a Roman harbour (Voorburg-Arentsburg) in the Netherlands. J
644 Archaeol Sci. 2014;41: 642-654. doi: 10.1016/j.jas.2013.09.009.
- 645 11. Fraiture P. Contribution of dendrochronology to understanding of wood
646 procurement sources for panel paintings in the former Southern Netherlands
647 from 1450 AD to 1650 AD. Dendrochronologia. 2009;27: 95-111. doi:
648 10.1016/j.dendro.2009.06.002.
- 649 12. Rich SA. Ship Timber as Symbol. Dendro-provenancing & Contextualizing
650 Ancient Cedar Ship Remains from the Eastern Mediterranean / Near East,
651 Ph.D. Thesis, KU Leuven. 2013.
- 652 13. Daly A, Nymoen P. The Bøle ship, Skien, Norway—Research History,
653 Dendrochronology and Provenance. Int. J. Naut. Archaeol. 2008;37: 153-170.
654 doi: 10.1111/j.1095-9270.2007.00157.x.

- 655 14. Bridge M. Locating the origins of wood resources: a review of
656 dendroprovenancing. *J Archaeol Sci.* 2012;39: 2828-2834. doi:
657 10.1016/j.jas.2012.04.028.
- 658 15. Gut U. Evaluating the key assumptions underlying dendro-provenancing: How to
659 spruce it up with a scissor plot. *Dendrochronologia.* 2018;52: 131-145. doi:
660 10.1016/j.dendro.2018.09.008.
- 661 16. Fowler A, Bridge M. Mining the British Isles oak tree-ring data set. Part A:
662 Rationale, data, software, and proof of concept. *Dendrochronologia.* 2015;35:
663 24-33. doi: 10.1016/j.dendro.2015.05.008.
- 664 17. Baillie MGL, Pilcher JR. A simple cross-dating program for tree-ring research.
665 *Tree-Ring Bull.* 1973;33: 7-14.
- 666 18. Hollstein E. Mitteleuropäische Eichenchronologie. Mainz am Rhein: Philipp von
667 Zabern; 1980.
- 668 19. Babst F, Bodesheim P, Charney N, Friend AD, Girardin MP, Klesse S, et al.
669 When tree rings go global: Challenges and opportunities for retro- and
670 prospective insight. *Quat Sci Rev.* 2018;197: 1-20. doi:
671 10.1016/j.quascirev.2018.07.009.
- 672 20. Schweingruber FH. Tree Rings and Environment: Dendroecology. Bern,
673 Stuttgart, Wien: Haupt; 1996.
- 674 21. Cook ER, Briffa KR, Meko DM, Graybill DA, Funkhouser G. The 'segment length
675 curse' in long tree-ring chronology development for palaeoclimatic studies.
676 *Holocene.* 1995;5: 229-237. doi: 10.1177/095968369500500211.

- 677 22. Fritts HC. Tree Rings and Climate. London, New York, San Francisco: Academic
678 Press; 1976.
- 679 23. Jansma E, Haneca K, Kosian M. A dendrochronological reassessment of three
680 Roman boats from Utrecht (the Netherlands): evidence of inland navigation
681 between the lower-Scheldt region in Gallia Belgica and the limes of Germania
682 inferior. *J Archaeol Sci.* 2014;50: 1-13. doi: 10.1016/j.jas.2014.07.019.
- 683 24. Grabner M, Buchinger G, Jeitler M. Stories about building history told by wooden
684 elements - case studies from Eastern Austria. *Int. J. Archit. Heritage.* 2018;12:
685 178-194. doi: 10.1080/15583058.2017.1372824.
- 686 25. R Core Team. R: A Language and Environment for Statistical Computing.
687 Vienna, Austria; 2019. Available: <https://www.R-project.org/>.
- 688 26. Cook ER, Kairiukstis LA, editors. Methods of Dendrochronology - Applications in
689 the Environmental Sciences. Dordrecht: Kluwer; 1990.
- 690 27. Cook ER. A Time Series Approach to Tree Ring Standardization. Ph.D. Thesis,
691 University of Arizona. 1985.
- 692 28. Stock JH, Watson MW. Introduction to Econometrics. 3rd ed. Boston: Pearson;
693 2015.
- 694 29. Bunn AG. A dendrochronology program library in R (dplR). *Dendrochronologia.*
695 2008;26: 115-124. doi: 10.1016/j.dendro.2008.01.002.
- 696 30. Akaike H. Information Theory and an Extension of the Maximum Likelihood
697 Principle. In: Parzen E, Tanabe K, Kitagawa G, editors. Selected Papers of
698 Hirotugu Akaike. New York, NY: Springer New York; 1998. pp. 199-213.

- 699 31. Ripley BD. Stochastic Simulation. Hoboken, NJ, USA: John Wiley & Sons, Inc;
700 1987.
- 701 32. Venables WN, Ripley BD. Modern Applied Statistics with S. 4th ed. New York,
702 NY: Springer; 2007.
- 703 33. Eissing, T. Holzversorgung, Holzverbrauch, Holzherkunft am Beispiel
704 Mitteldeutschlands. In: Dix A, Schenk W, editors. Konsum und
705 Kulturlandschaft. Bonn: ARKUM e.V.; 2010. pp. 87-106.
- 706 34. Abanda A, Mori U, Lozano JA. A review on distance based time series
707 classification. Data Min Knowl Discov. 2019;33: 378-412. doi:
708 10.1007/s10618-018-0596-4.
- 709 35. Kaya H, Gündüz-Öğüdücü Ş. A distance based time series classification
710 framework. Inf Syst. 2015;51: 27-42. doi: 10.1016/j.is.2015.02.005.
- 711 36. Xing Z, Pei J, Keogh E. A brief survey on sequence classification. SIGKDD
712 Explor. Newsl. 2010;12: 40-48. doi: 10.1145/1882471.1882478.
- 713 37. Pearson K. Note on regression and inheritance in the case of two parents.
714 Proceedings of the Royal Society of London. 1895;58: 240-242.
- 715 38. Parker RE. Introductory Statistics for Biology. 2nd ed. Cambridge: Cambridge
716 University Press; 1979.
- 717 39. Eissing T. Zum Problem der Interpretation dendrochronologischer Datierung bei
718 Flossholz am Beispiel Bamberg. In: Arbeitskreis für Hausforschung, editor.
719 Spuren der Nutzung in historischen Bauten. Marburg: Jonas; 2007. pp. 23–36.

- 720 40. Daly A. The Karschau Ship, Schleswig-Holstein: Dendrochronological Results
721 and Timber Provenance. *Int J Nautical Archaeology*. 2007;36: 155-166. doi:
722 10.1111/j.1095-9270.2006.00103.x.
- 723 41. Daly A. Timber, Trade and Tree-rings. A dendrochronological analysis of
724 structural oak timber in Northern Europe, c. AD 1000 to c. AD 1650. Ph.D.
725 Thesis, University of Southern Denmark. 2007.
- 726 42. Rodríguez-Trobajo E, Domínguez-Delmás M. Swedish oak, planks and panels:
727 dendroarchaeological investigations on the 16th century *Evangelistas*
728 altarpiece at Seville Cathedral (Spain). *J Archaeol Sci*. 2015;54: 148-161. doi:
729 10.1016/j.jas.2014.11.039.
- 730 43. Wazny T. Baltic timber in Western Europe – an exciting dendrochronological
731 question. *Dendrochronologia*. 2002;20: 313-320. doi: 10.1078/1125-7865-
732 00024.
- 733 44. Karanitsch-Ackerl S, Holawe F, Laaha G, Wimmer R, Grabner M. Parameter-
734 specific hydroclimatic sensitivity of a low-elevation network of living and
735 historical tree-ring series from north-eastern Austria. *Dendrochronologia*.
736 2017;45: 39-51. doi: 10.1016/j.dendro.2017.06.004.
- 737 45. Hellmann L, Tegel W, Geyer J, Kirdyanov AV, Nikolaev AN, Eggertsson Ó, et al.
738 Dendro-provenancing of Arctic driftwood. *Quat Sci Rev*. 2017;162: 1-11. doi:
739 10.1016/j.quascirev.2017.02.025.
- 740 46. Fowler AM, Bridge MC. Empirically-determined statistical significance of the
741 Baillie and Pilcher (1973) t statistic for British Isles oak. *Dendrochronologia*.
742 2017;42: 51-55. doi: 10.1016/j.dendro.2016.12.006.

- 743 47. Fowler AM, Bridge MC, Boswijk G. An empirical resampling method for
744 determining optimal high-pass filters used in correlation-based tree-ring
745 crossdating. *Dendrochronologia*. 2017;44: 84–93. doi:
746 10.1016/j.dendro.2017.04.003.
- 747 48. Bridge M. Can dendrochronology be used to indicate the source of oak within
748 Britain? *Vernacular Architecture*. 2000;31: 67–72.
749 doi:10.1179/vea.2000.31.1.67
- 750 49. García-González I. Comparison of Different Distance Measures for Cluster
751 Analysis of Tree-Ring Series. *Tree Ring Res.* 2008;64: 27–37.
752 doi:10.3959/2007-2.1
- 753 50. Haneca K, Wazny T, van Acker J, Beeckman H. Provenancing baltic timber from
754 art historical objects: Success and limitations. *J Archaeol Sci.* 2005;32: 261–
755 271. doi:10.1016/j.jas.2004.09.005
- 756 51. Savva Y, Oleksyn J, Reich PB, Tjoelker MG, Vaganov EA, Modrzynski J.
757 Interannual growth response of Norway spruce to climate along an altitudinal
758 gradient in the Tatra Mountains, Poland. *Trees*. 2006;20: 735–746.
759 doi:10.1007/s00468-006-0088-9
- 760 52. Wilson R, Hopfmüller M. Dendrochronological investigations of Norway spruce
761 along an elevational transect in the Bavarian Forest, Germany.
762 *Dendrochronologia*. 2001;19: 67–79.
- 763 53. Frank D, Esper J. Characterization and climate response patterns of a high-
764 elevation, multi-species tree-ring network in the European Alps.
765 *Dendrochronologia*. 2005;22: 107–121. doi:10.1016/j.dendro.2005.02.004

- 766 54. Dittmar C, Eissing T, Rothe A. Elevation-specific tree-ring chronologies of
767 Norway spruce and Silver fir in Southern Germany. *Dendrochronologia*.
768 2012;30: 73–83. doi:10.1016/j.dendro.2011.01.013
- 769 55. King GM, Gugerli F, Fonti P, Frank DC. Tree growth response along an
770 elevational gradient: Climate or genetics? *Oecologia*. 2013;173: 1587–1600.
771 doi:10.1007/s00442-013-2696-6
- 772 56. Kolář T, Čermák P, Trnka M, Žid T, Rybníček M. Temporal changes in the climate
773 sensitivity of Norway spruce and European beech along an elevation gradient
774 in Central Europe. *Agric For Meteorol*. 2017;239: 24–33.
775 doi:10.1016/j.agrformet.2017.02.028
- 776 57. Lyu L, Suvanto S, Nöjd P, Henttonen HM, Mäkinen H, Zhang Q-B. Tree growth
777 and its climate signal along latitudinal and altitudinal gradients: Comparison of
778 tree rings between Finland and the Tibetan Plateau. *Biogeosciences*. 2017;14:
779 3083–3095. doi:10.5194/bg-14-3083-2017
- 780 58. König J, Günther B, Bues CT. Application of multivariate cross-dating to historical
781 timbers with less than 50 tree rings from the Albrechtsburg Castle and the
782 Meißen Cathedral, Saxony. In: Elferts D, Brumelis G, Gärtner H, Helle G,
783 Schleser G, editors. *Tree Rings in Archeology, Climatology and Ecology*;
784 2007.
- 785 59. Wilson R, Wilson D, Rydval M, Crone A, Büntgen U, Clark S, et al. Facilitating
786 tree-ring dating of historic conifer timbers using Blue Intensity. *J Archaeol Sci*.
787 2017;78: 99–111. doi: 10.1016/j.jas.2016.11.011.

788 **9 Supporting information**

789 **S1 R package: *ad hoc* pseudo.series package.**

790 **S2 Figs: Additional Figures 1-9.**

791 **S3 Tables: Additional Tables 1-9.**

792 **S4 Text: Additional details on prefs established with poc approaches.**

793 **S5 Figs: Additional Figures 10-13.**