RESEARCH PAPER

The International Tree-Ring Data Bank (ITRDB) revisited: Data availability and global ecological representativity

Shoudong Zhao[1] | Neil Pederson[2] | Loïc D'Orangeville[2,3] | Janneke HilleRisLambers[4] | Emery Boose[2] | Caterina Penone[5] | Bruce Bauer[6] | Yuan Jiang[1] | Rubén D. Manzanedo[2,4] (iD)

[1]State Key Laboratory of Earth Surface Processes and Resource Ecology, Faculty of Geographical Science, Beijing Normal University, Beijing, China

[2]Harvard Forest, Harvard University, Petersham, Massachusetts

[3]Faculty of Forestry and Environmental Management, University of New Brunswick, Fredericton, Canada

[4]Biology Department, University of Washington, Seattle, Washington

[5]Institute of Plant Sciences, University of Bern, Bern, Switzerland

[6]National Centers for Environmental Information, National Oceanic and Atmospheric Administration, Asheville, North Carolina

Correspondence
Rubén D. Manzanedo. Harvard Forest, Harvard University. 324 N Main St Petersham MA 01366.
Email: rdmanzanedo@fas.harvard.edu

Funding information
National Natural Science Foundation of China, Grant/Award Number: NSFC 41630750; Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung, Grant/Award Number: Early Post.doc Mobility grant # P2BEP3_175293; National Science Foundation, Grant/Award Number: PalEON MacroSystems Biology grant #1241930

Editor: Alexandre Antonelli

Abstract

Aim: The International Tree-Ring Data Bank (ITRDB) is the most comprehensive database of tree growth. To evaluate its usefulness and improve its accessibility to the broad scientific community, we aimed to: (a) quantify its biases, (b) assess how well it represents global forests, (c) develop tools to identify priority areas to improve its representativity, and d) make available the corrected database.

Location: Worldwide.

Time period: Contributed datasets between 1974 and 2017.

Major taxa studied: Trees.

Methods: We identified and corrected formatting issues in all individual datasets of the ITRDB. We then calculated the representativity of the ITRDB with respect to species, spatial coverage, climatic regions, elevations, need for data update, climatic limitations on growth, vascular plant diversity, and associated animal diversity. We combined these metrics into a global Priority Sampling Index (PSI) to highlight ways to improve ITRDB representativity.

Results: Our refined dataset provides access to a network of >52 million growth data points worldwide. We found, however, that the database is dominated by trees from forests with low diversity, in semi-arid climates, coniferous species, and in western North America. Conifers represented 81% of the ITRDB and even in well-sampled areas, broadleaves were poorly represented. Our PSI stressed the need to increase the database diversity in terms of broadleaf species and identified poorly represented regions that require scientific attention. Great gains will be made by increasing research and data sharing in African, Asian, and South American forests.

Main conclusions: The extensive data and coverage of the ITRDB show great promise to address macroecological questions. To achieve this, however, we have to overcome the significant gaps in the representativity of the ITRDB. A strategic and organized group effort is required, and we hope the tools and data provided here can guide the efforts to improve this invaluable database.

KEYWORDS
bias analysis, big data, data accessibility, Dendrochronology, dendroecology, meta-analysis, tree growth, tree-ring research

1 | INTRODUCTION

Understanding terrestrial ecosystems at global scales is a fundamental challenge that demands datasets with large temporal and spatial coverage (Babst, Poulter, Bodesheim, Mahecha, & Frank, 2017; Blume-Werry, Kreyling, Laudon, & Milbau, 2016). Dendrochronological data (i.e., data from tree-ring measurements) have proven to be an important resource, providing valuable information on large-scale palaeoclimatic (e.g., Fritts, Blasing, Hayden, & Kutzbach, 1971; Graumlich & Brubaker,1986; Jacoby & D'Arrigo, 1989), ecological (e.g., Charney et al., 2016; Foster, D'Amato, & Bradford, 2014; Fritts & Swetnam, 1989; Jucker, Bouriaud, Avacaritei, & Coomes, 2014; Pederson et al., 2014), and geomorphological (e.g., Stoffel & Bollschweiler, 2008), and climatic (e.g., Cook, Woodhouse, Eakin, Meko, & Stahle, 2004; Schweingruber, Bräker, & Schär, 1979; Stahle, Cleaveland, & Hehr, 1985; Stockton & Jacoby, 1976) processes with a spatiotemporal coverage and resolution difficult to match by another approaches (e.g., Babst et al., 2017). While new techniques, such as remote sensing or LiDAR, grant access to information at even larger scales, tree-ring data offer a long-term perspective more representative of the time-scale of forest dynamics, providing a closer mechanistic insight on how trees respond to environmental change. The scarcity of long-term data is a long-standing problem in ecology, particularly forest ecology (Wiens, 1989). The extreme longevity of trees makes it virtually impossible to encompass the complete life of even a single individual without using tree-ring methods. A comprehensive and global database of tree-ring growth, such as the International Tree-Ring Data Bank (ITRDB), can contribute to earth-system sciences by incorporating this unique spatiotemporal coverage (Babst et al., 2017). To achieve this, however, we first need to understand the strengths and weaknesses of the current tree-ring data collections.

Established for the preservation of high-quality dendrochronological data in 1974 (Grissino-Mayer & Fritts, 1997), the ITRDB is still the primary archive of global tree-ring data. It is a repository for tree-ring measurements (ring width and density data) and other wood anatomical features (earlywood-latewood measurements or isotopic composition). The ITRDB archives chronologies (standardized mean time series) for over 4,000 locations of 226 tree species over all continents (except for treeless Antarctica), spanning from 6000 BCE to present dates. The ITRDB is open for users to download and use its data with the only requirement of citing data contributors. Despite its potential, only a handful of studies have used significant portions of the dataset to explore global processes. These works have also pointed at the weaknesses in the current dataset that challenged their ability to interpret their results (Table 1; Charney et al., 2016; Gedalof & Berg, 2010; St. George, Ault, & Torbenson, 2013; Mina, Martin-Benito, Bugmann, & Cailleret, 2016; Tei et al., 2017; Williams et al., 2013). As a result, and despite recent efforts, the ITRDB remains underutilized in comparison with other global databases in ecology, e.g., TRY (www.try-db. org).

One reason for the limited use of the ITRDB is its accessibility and data formatting. The ITRDB can only be downloaded as a collection of independent files in the Tucson format. Developed in the 1960's by Hal Fritts, the Tucson format was designed around the 80-column Hollerith computer punch cards method of inputting data to computers. It included innovative features such as delimiters to signal measurement precision, sample identity information to avoid pseudo-replication in further analyses, and metadata headers. The approach effectively introduced computing to tree-ring science and revolutionized dendrochronological methods (E. Cook *pers. commun. Emails, Jan. 2–3, 2018*). However, it is now generally not used outside of dendrochronology due to its inflexibility to accommodate

TABLE 1 Literature examples using ITRDB in diverse disciplines, which have also mentioned the bias in the existing dendrochronological data that reduced their potential to draw more general conclusions

Discipline	Ecological mechanism	Bias discussed	Priority sampling	Reference
Tree Physiology	Locally absent rings	Bias towards continental United States and western Europe	Africa, Middle East, India, and eastern Asia	St. George et al., 2013
Ecology	CO_2 fertilization induced growth increase	Bias towards sites where growth is sensitive to environmental variability	Suitable environment for trees	Gedalof & Berg, 2010
	Forest response to increasing aridity and warmth			Williams et al., 2010
	Forest response to increased water-use efficiency			Charney et al., 2016
	Forest response to drought	Bias towards gymnosperms and evergreen tree species	Angiosperms and broad-leaved tree species	Vicente-Serrano, Camarero, & Azorin-Molina, 2014
	Volcanic eruptions and net primary production	Bias towards dominant trees	Understory trees	Krakauer & Randerson, 2003
		Bias towards temperate forests	Tropical forests	
Climatology	Temperature reconstruction at hemispheric scales	Need to be updated to present	Globally update	Wilson et al., 2007

extra data. Transforming the ITRDB to a commonly used and versatile format has become a necessary step to ensure its use by the broader scientific community.

Additionally, the limited financial support to ITRDB means that the quality of each contributed dataset might not be completely checked and verified. This has resulted in an accumulation of formatting errors, unreadable files, and empty data files. While some efforts have been made to improve the accessibility of ITRDB data for large-scale research (Breitenmoser, Brönnimann, & Frank, 2014; Zang, 2015), a full reformatting, harmonizing, and republication of the ITRDB is deeply needed. Such a step would globally save hours of redundant work, increase the consistency in the results, and improve the use of the database.

The ITRDB is formed by voluntary and generous contributions of hundreds of scientists worldwide and, consequently, reflects their research priorities and interests, including those of various funding agencies. Since its origins, tree-ring research has primarily focused on reconstructing climate, investigating global warming, and studying growth–climate correlations in dominant trees. These studies followed a sampling design that increases the climate signal (at the cost of reducing the strength of other ecological signals, such as competition, diversity, or management) and the spatial coverage (at the cost of reducing the number of sampled trees per location, which frequently include as few as 10–20 individuals) (discussed in Sullivan & Csank, 2016). Developers of the ITRDB recognized these biases and planned to expand the database to accommodate tree-ring data collected for a wider variety of scientific questions (Grissino-Mayer & Fritts, 1997). However, this never came to full fruition, and climate-focused studies still dominate the ITRDB.

Dendroecology, or ecologically focused tree-ring research, is quickly gaining momentum and, in so doing, stirring a re-evaluation of classical dendrochronological methods (Amoroso, Daniels, Baker, & Camarero, 2017). The low sample size, tree selection criteria, reduced information on the microhabitat conditions, and forest structure used in many dendrochronological studies are increasingly questioned for the study of forest dynamics and ecological processes (Brienen, Gloor, & Ziv, 2017; Briffa & Melvin, 2011; Davis, Hessl, Scott, Adams, & Thomas, 2009; Nehrbass-Ahles et al., 2014; Pederson et al., 2014; Sullivan & Csank, 2016). However, the current ITRDB cannot accommodate the extra metadata demanded by more ecological studies (Foster et al., 2014). Similarly, the temporal coverage (Babst et al., 2017) and geographical biases (Amoroso et al., 2017) in the database raise concerns on its representativity. A quantification of biases in the ITRDB and its appropriateness to study global ecological processes is urgent to inform future sampling and gives an adequate context to interpret past and future results from the ITRDB.

To explore these issues, we addressed three linked goals. First, we reformatted, harmonized, and combined all the available data in the ITRDB into a new easy-to-access dataset in the R programming language. Second, we quantified its spatial, climatic, temporal, taxonomic, and ecological biases and its representation of global ecosystems. Finally, we developed a user-defined priority sampling index (PSI) that

highlights areas that would improve ITRDB's representativity while trying to match the researcher's scientific priorities in terms of the main environmental conditions they focus on. We highlight the main results provided by this index and discuss its implications for future research.

2 | MATERIALS AND METHODS

2.1 | ITRDB download and data processing

The ITRDB database is available via the repositories of the National Climatic Data Center (https://www1.ncdc.noaa.gov/pub/data/paleo/treering). We used the last available version, 7.13, as of June 2017 (last accessed October 12th, 2017). We downloaded 8326 individual files in Tucson format containing information on tree-ring width, wood characteristics, or isotopic composition. Files were uploaded to R (R Core Team, 2017) using the dplR package (version 1.6.6) (Bunn, 2008). We identified conflictive files using the read.rwl function of dplR. Error outputs were collected and automatically classified into error types to aid further corrections. We integrated the available metadata for each location (species, elevation, spatial coordinates, first and last year of data, and contributor) into a single file. A full description of the encountered issues, correction strategies, modifications, and metadata can be found in Appendix 1 for full traceability. The final, compiled, and harmonized database is available via https://www.ncdc.noaa.gov/paleo/study/25570. All analyses were performed in R (version 3.4.3) and can be replicated using the available R script (https://doi.org/10.5061/dryad.kh0qh06).

2.2 | Determining the representativity of the ITRDB

We calculated the coverage and representation of ITRDB sites regarding important environmental variables. We combined site metadata with publicly available information of key environmental variables (Table 2). See Table S1 for the importance associated with a representative database for each of the selected variables.

First, to explore the ITRDB's spatial representation, we calculated the spatial density of sites using a kernel density estimation map with a 10° bandwidth value. To quantify the spatial bias, we compared the number of chronologies sampled in each continent with what would be expected by random chance, given the continent's geographical size (see Gonzalez et al., 2016). The confidence intervals around the null distribution were calculated using 999 bootstrapped random distributions of the 3621 unique sampling locations in the ITRDB (excluding Antarctica). Sampling frequencies outside of the range of bootstrapped values were interpreted as significantly over- or underrepresented in the database. By "overrepresented", we do not mean to imply that further scientific research may not be needed in those ecosystems. Representativity here refers to an unbalance in the sampling efforts between regions, species, habitats, etcetera. These unbalances can influence the conclusions drawn from the database, which would be strongly shaped by the response of those overrepresented regions. It should be understood,

TABLE 2 Description and origin of the environmental variables used in the bias analyses. When variables were combined, they were resampled to a common 2° raster grid resolution

Variable	Code	Description	Original resolution	Reference	Source	Accessed
Continent	C	Continent of the location (7 continents model)	Continent	—	Calculated from ITRDB sampling coordinates	12.10.17
Temperature limitation on growth	TI	Temperature limitation on vegetation productivity	0.5° raster grid	Churkina & Running, 1998	https://databasin.org/datasets/0b13ef1088b645a29ee49f8aaaabb44	31.10.17
Water limitation on growth	WI	Water limitation on vegetation productivity	0.5° raster grid	Churkina & Running, 1998	https://databasin.org/datasets/321550b8dfeb459ca5856508621bb1bc	31.10.17
Elevation	EI	Geographical elevation above sea level (m)	2.5' raster grid	Hijmans, Cameron, Parra, Jones, & Jarvis, 2005	Derived from ITRDB metadata. Complemented with WorldClim DEM http://www.worldclim.org/formats1	30.10.17
Climatic regions	CI	Main zones from Köppen-Geiger climatic classification	Variable, polygons	Kottek, Grieser, Beck, Rudolf, & Rubel, 2006	http://koeppen-geiger.vu-wien.ac.at/present.htm	30.10.17
Vascular plant diversity	Pdiv	Number of plant species by terrestrial ecoregion	Variable, polygons	Kier et al., 2005	https://databasin.org/datasets/43478f840ac84173979b22631c2ed672	6.11.17
Associated forest diversity	Assdiv	Total added number of species of birds, mammals, and amphibians	10 Km raster grid	Various	Calculated from taxa maps downloaded at: http://biodiversitymapping.org/wordpress/index.php/download/	6.11.17
Need to update	U	The difference between 2017 and the most recent year with tree-ring growth data in a 2° × 2° sampling grid worldwide	2° raster grid	—	Calculated from ITRDB sampling coordinates and data	12.10.17
Seasonality	Seas	Locations worldwide with high seasonality, thus likely to contain tree-ring forming species	10' raster grid	Hijmans et al., 2005	Derived from ITRDB metadata and temperature seasonality maps by Worldclim (BIO4) http://www.worldclim.org/bioclim	06.11.17
Vegetation cover	Vcov	Vegetation lower than 1 m was excluded to remove low grasslands and bare ground areas	1 Km raster grid	Simard, Pinto, Fisher, & Baccini, 2011	Global vegetation map https://landscape.jpl.nasa.gov/	06.11.17

therefore, as a call for doubling our efforts to study poorly represented areas and not to reduce that of the better studied ones.

To calculate relative frequencies per species and taxonomic group, we considered locations with more than one measurement on the same samples as unique sampling sites (e.g., isotopic composition and ring width measurements). Datasets from different species with same coordinates were considered as two different species samples for the taxonomical analysis, but as one sampling location for all further analyses. Independent site estimations of growth limiting factors (water, temperature) were obtained from publicly available raster data (see Table 2). These factors were based on Churkina and Running (1998), who assessed the relative importance of several climatic controls on terrestrial biome net primary productivity. We did not include radiation limitation, as the original paper found a very small proportion of the global biomes limited by radiation. We calculated frequency density functions of sampled sites for both water and temperature limitation on growth (Figure S1). Similarly, we explored the distribution of ITRDB locations between main climatic regions using the climatic classification by Köppen-Geiger and site coordinates (Figure S2).

The distribution of sampled elevations (Figure S3) was calculated from publicly available digital elevation models (Table 2), after checking for consistency with ITRDB metadata. To correct for the large imbalance in global area between high and low elevations, we standardized the frequencies of ITRDB sites using a global frequency distribution of elevations. Consequently, negative (positive) values of this ratio show areas with lower (higher) number of chronologies than expected, given the global proportion of land with that elevation.

We similarly calculated the vascular plant diversity and associated animal forest biodiversity captured by ITRDB sites (Figure S4). Vascular plant diversity was calculated as the total number of plant species by terrestrial biome (Kier et al., 2005), while the associated forest biodiversity was the sum of the species diversity of amphibians, birds, and mammals for each sampling location (Table 2). Unfortunately, no richness data with global resolution are currently available for other important forest taxa such as arthropods, soil bacteria, or fungi. Should this information become available, it could easily be implemented in our code.

Finally, we determined the extent to which existing datasets include information relative to recent climate change. Specifically, we calculated the most recent year with tree-ring data for a 2° × 2° grid worldwide (WGS84 projection). Grid-cells were assigned an update priority ranging from 0, for grid-cells last sampled in 2017; to 1, for grid-cells last sampled more than 50 years ago. Since almost all ITRDB samples were collected after 1967 (see Figure S5 and Babst et al., 2017), setting the maximum index value for the few medieval or prehistoric samples would have resulted in a clump of very low values for most of the remaining sites, rendering differences between datasets meaningless. In addition, the last 50 years represent a time of accelerated climate change. It is important to include as many chronologies as possible to study this critical period. Grid-cells with zero sampling locations were considered as NA, by definition.

To further investigate taxonomic biases, we compared the relative abundance of species within the ITRDB to the current forest composition. This was conducted for 30 Eastern US states, for which we had access to precise inventory data from the most recent USDA Forest Inventory and Analysis Program. The relative abundance of each species in the ITRDB was computed as its sampling frequency in plots from across the region (see map in Figure S6). Current forest composition was computed as species relative basal area for trees with diameter at breast height (DBH) ≥20 cm across the region. This approach was also used to assess the proportion of conifer versus broadleaf species in the region.

2.3 | Priority sampling index

We calculated an independent priority index for each environmental variable using their distribution of values or density frequency curve. These indices ranged from 0 to 1, where 0 are well represented in the ITRDB, and 1 are environmental conditions covered by ≤1 locations, and thus greatly underrepresented. The response functions were then extrapolated to the original variables global raster to map areas with underrepresented characteristics worldwide. These indices indicate the potential of a given location to improve the representativity of the ITRDB for a given factor. Response functions were calculated for continents, climatic regions, elevations, need for series update, water and temperature limitations on growth, vascular plant diversity, and associated forest diversity.

Finally, individual response functions were combined into a single global priority sampling index (PSI) that indicates the potential of any given grid-cell to improve the ITRDB with respect to all considered environmental variables. PSI was calculated as a weighted mean of all priority indices from each 2° grid-cell as:

$$PSI_i = \frac{\sum_1^n w_i X_i}{\sum_1^n w_i}$$

where w_i represents the weight given to each environmental variable X_i (see Table 2 and Table 3). Since the weight assigned to each variable depends on subjective research priorities, the PSI calculation tool, provided in the supplementary scripts, allows researchers to define their own weighting schemes to reflect their own research priorities while keeping into consideration how their work can contribute to a more comprehensive ITRDB.

To illustrate the usefulness of PSI, we defined weights for three basic scenarios, based on the authors' expertise (Table 3). The scenarios we defined are: equal weights to all considered factors (*Equal Weights*), priority for dendroecological and biodiversity research

TABLE 3 Weight scores in the priority sampling indices maps shown in Figure 5. Variable code as in Table 2

Scenario	w_C	w_{Cl}	w_{El}	w_U	w_{Tl}	w_{Wl}	w_{Pdiv}	w_{Assdiv}
Equal Weights	1	1	1	1	1	1	1	1
Ecological Research	1	1	0.5	0.1	0.1	0.1	1	1
Climatic Research	1	0.1	0.1	1	0.5	0.5	0.1	0.1

(*Ecological Research*), and focus on climate reconstruction or climate sensitivity research (*Climatic Research*). The *Ecological Research* scenario prioritized greater continental and ecoregional coverage of the ITRDB, giving also high priority to places that would increase the gradient of diversities covered in the database. The *Climatic Research* scenario also prioritized the continental coverage to improve the interpolation of climatic reconstructions but assigned a high priority to updating old series. Updating old series can be an easy and effective way to improve spatial and temporal coverage of climatic reconstructions, since these locations are certain to have crossdating species, a fairly well-identified climatic signal, and in some cases, available meteorological data, facilitating dendrochronological analyses. Covering a broad range of water and temperature limitations was also considered important for climatic reconstructions to capture the effects of climate change on different types of vegetation.

The resulting priority maps can be filtered by two important variables for dendrochronology, temperature seasonality, and forest coverage. These filters are optional, and we present here both filtered and unfiltered results. Highly seasonal areas are more likely to contain ring forming species, as seasonality provides the physiological constraints that trigger the formation of visible growth rings. While recent tropical studies show this is not necessarily the case, tropical dendrochronology is still in the early stages of development and still challenging. The seasonality filter removes areas with lower temperature seasonality than that of the first quartile distribution of the complete ITRDB (see O'Donnel & Ignizio, 2012 for more information and Table 2 for data sources). Taxonomical differences should also be considered. For example, Australia has 2,400+ woody species, according to the Dadswell wood collection, but only 12 of them have been so far reported to be useful for dendrochronological purposes (Heinrich & Allen, 2013); and attempts to develop

chronologies for other species have fallen short (P. Baker *pers. commun.* Nov 18th, 2017). The forest coverage filter removes areas with vegetation ≤1 m in height, which are less likely to have woody species useful for dendrochronology. Researchers focusing on creeping trees (e.g., *Salix arctica* in Schmidt, Baittinger, Kollmann, & Forchhammer, 2010), or shrubs (e.g., Myers-Smith et al., 2011, and www.ShrubHub.biology.ualberta.ca), would likely disable this filter. Vegetation height was calculated from Jet Propulsion Laboratory of NASA (see Table 2).

3 | RESULTS

3.1 | Formatting issues and their distribution

Of the 8326 individual ITRDB files, 28.8% had formatting problems, 20.6% of which prevented the file from being read by dplR's reading functions and 8.2% had minor issues that could be addressed by dplR's reading functions. Formatting errors were not limited to any geographical region, although European, South American, and African datasets had the highest proportion of conflictive files (Figure 1). Many formatting errors were related to the use of non-standard Latin characters such as accents or umlauts that conflict with UTF-8 encoding.

A more challenging issue is that of repeated cores (i.e., data duplication). In some files, cores had the same identification number but different (and sometimes overlapping) measurements. These conflictive cores represent a small fraction of the database (<1% of the cores) but pose a challenge for researchers needing to distinguish individual cores. We removed those repeated cores with identical growth values (fully repeated cores) or with overlapping and conflicting values. When the cores did not conflict or were consistent in their overlapping values, we considered they were segment
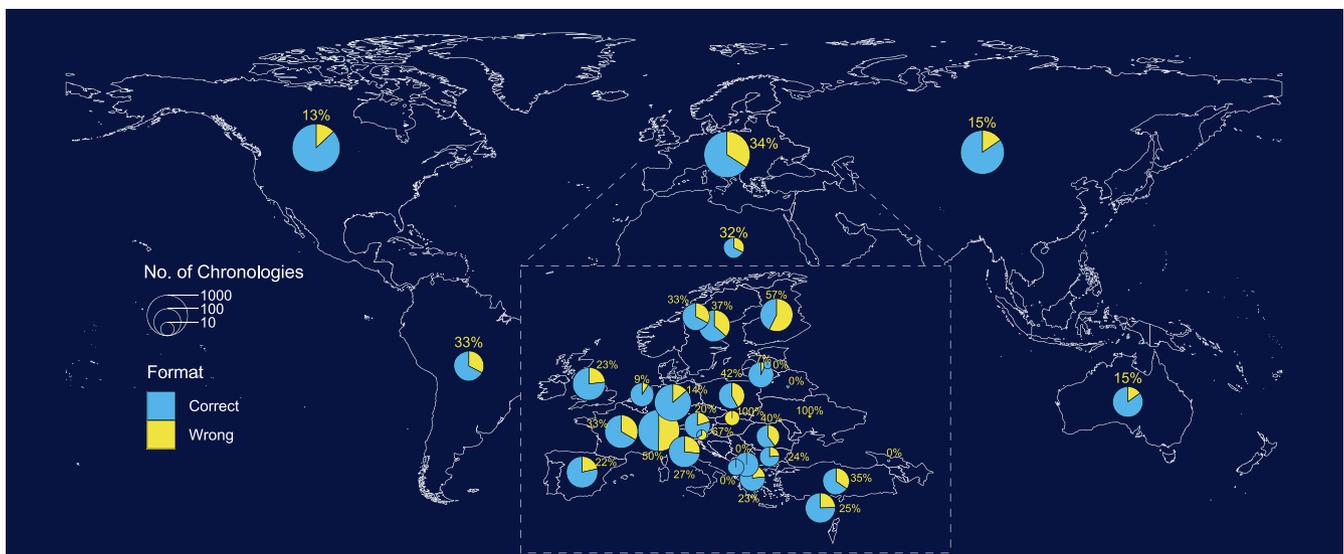
FIGURE 1 Formatting issues in the International Tree-Ring Databank (ITRDB) were globally distributed, though the highest proportion was found in European studies (inset map). Pie chart size represents the number of chronologies from each area in the database. The proportion of datasets with formatting issues (yellow) and correctly formatted (blue) is shown for each region or country. The proportion of datasets with formatting issues is also shown in yellow text

FIGURE 2 Spatial representation of the International Tree-Ring Data Bank (ITRDB). The density of individual tree-ring chronologies (yellow dots) shows the stronger coverage of the Northern Hemisphere, particularly of western North America. The density kernel was calculated with a 10° latitudinal bandwidth with brighter colors showing areas with higher data coverage. Repeated chronologies per coordinates, due to multiple measurements on the same samples or multispecies records, were considered as unique sampling locations

measurements of the same core and added a core label to them (see 'error_
correction_log.csv' in Appendix 1).

3.2 | Spatial, species, and limiting factors representativity

We found a large imbalance in the global distribution of the ITRDB (Figure 2). ITRDB collections were clustered in western North America, particularly in the rain shadow of the Rocky Mountains, and in western Europe, particularly over the Swiss Alps. Smaller clusters dispersed across the world seemed linked to important tree-ring institutions or conifer-dominated areas, such as Patagonia in South America, northern Siberia, and the Tibetan Plateau in central Asia. Some clusters coincided with historical sampling collections, for example New Zealand, as well as from emerging dendrochronological areas in Oceania. Our analysis clearly stressed, though, the lack of tree-ring data in large portions of South America, Oceania, and Asia;
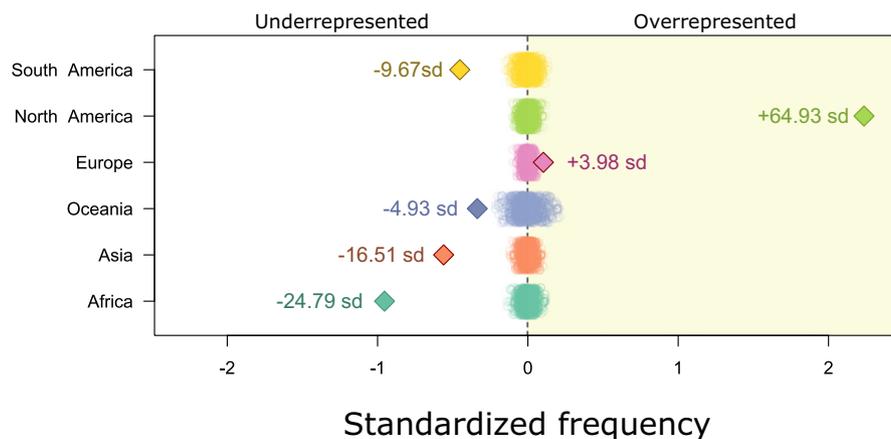
FIGURE 3 Quantification of the spatial bias in the ITRDB. The number of chronologies per continent (y-axis) is standardized according to the mean expected frequency per continent based on a null geographical model. Dots around 0 show the results of this null model (999 iterations of randomly located sampling sites with same sample size as the ITRDB). The ITRDB scores (diamonds) show the underrepresentation of South American, Oceanian, Asian, and African forests in the ITRDB database and the large overrepresentation of North American forest. The distance (measured in standard deviations from the mean of each distribution) is also displayed
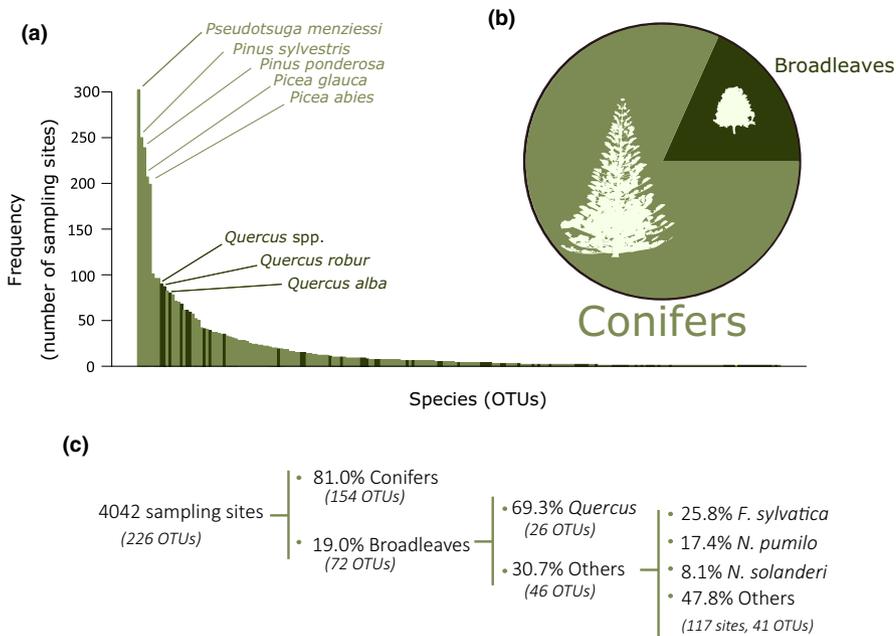
FIGURE 4 Species bias in the ITRDB. **a**, Relative species frequencies for each Operational Taxonomic Unit (OTUs) represented in the ITRDB. Most frequent conifer and broadleaf species are emphasized. (b) The ITRDB is sharply dominated by conifer species (81.7%) compared with broadleaves (18.3%). (c) Furthermore, the partition of the broadleaves group in the ITRDB show the high number of chronologies related to species of the *Quercus* genus, while other widespread species, such as *Fagus sylvatica*, are represented by a reduced number of chronologies

the scarcity of data from the Russian tundra, and an almost complete lack of information on the growth of African trees.

North America had almost 65 standard deviations, greater number of sampling sites than expected by random chance (Figure 3). It is important to notice, however, that not all areas within each continent were equally sampled (Figure 2). For example, while North America is strongly overrepresented, Canada and Mexico have lower sampling density than the US, and similarly, within the US, not all areas were equally represented. All the other continents, save Europe, were underrepresented compared with their geographical size. The sampling intensity in Europe is slightly above the variability

expected by random chance (+3.98 *SD*). A large proportion of sampling sites were located in warm and cold temperate regions (Figure S2), while areas with equatorial climate were only marginally represented.

Only the distribution of sampling sites by elevation was fairly representative of the global distributions (Figure S3), although areas under 1,000 masl and above 4,500 masl were still slightly underrepresented, and medium to high elevations (up to 4,000 masl) slightly overrepresented.

The ITRDB captures a range of temperature- and water-limited regions (Figure S1). Sites where water is expected to be the main
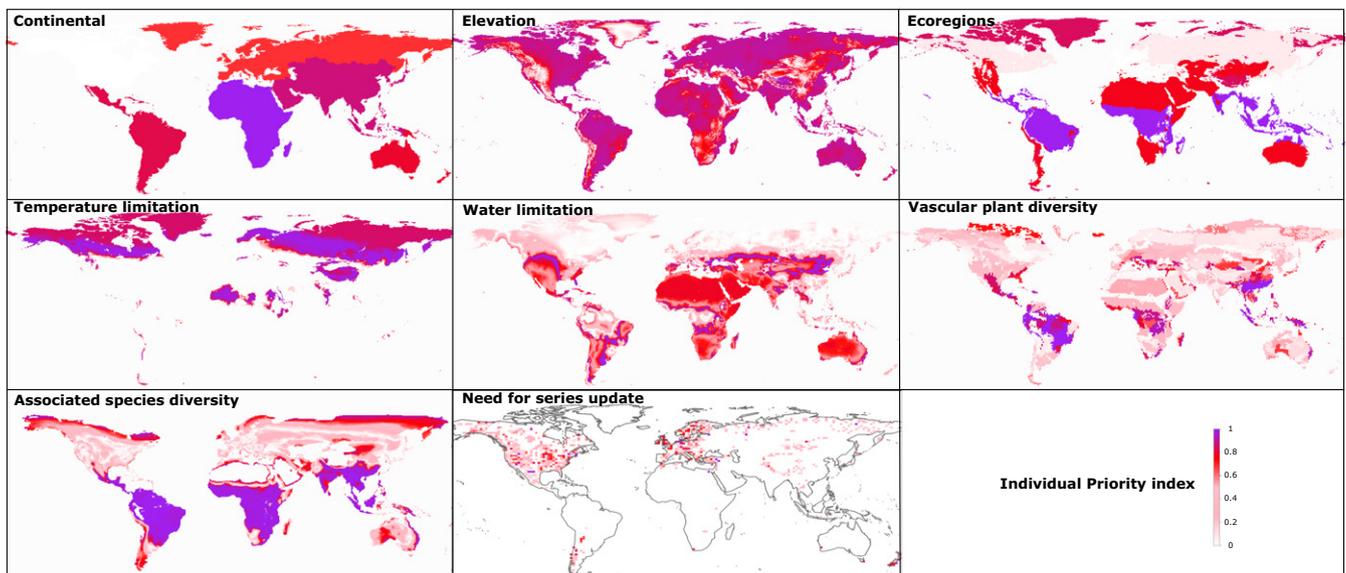
FIGURE 5 Extrapolated priority values for each individual variable that composes the global priority sampling index (PSI), that is shown in Fig. 6. Each index ranges between 0 and 1, where 1 are places with the highest probability to reduce the bias in the ITRDB in regards of that variable. Values of 0 represent areas with the already well represented areas in the ITRDB in regard to that variable

Unfiltered   ## Filtered
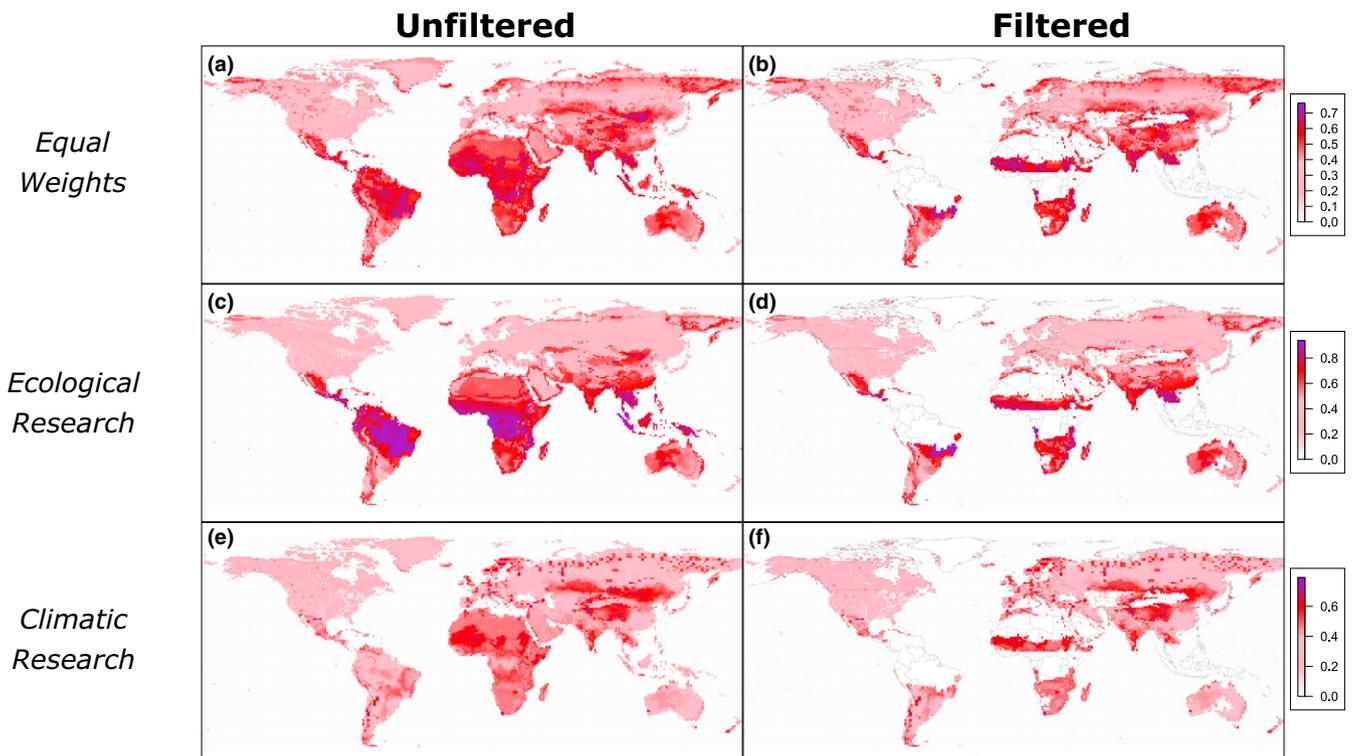
FIGURE 6 Priority sampling scenarios. The priority sampling index (PSI) shows the overall potential contribution of a given location to improve the data coverage of the ITRDB on regards of continents, climatic regions, elevations, temperature limitation on growth, water limitation on growth, need for series update, vascular plant diversity, and associated forest diversity (i.e. sum of amphibian, bird, and mammal species). Three scenarios with different variable importance (see Table 3) are shown: *Equal weights* (a,b), *Ecological research* (c,d) and *Climatic research* (e,f). Filtered maps (b,d,f) show only areas with high seasonality and vegetation higher than 1m, thus likely to have species that form distinctive tree-rings. Administrative regions are shown in the filtered maps for context

limiting factor were located predominantly in western North America, the Mediterranean Coast, and central Asia. By contrast, temperature limitation was captured by the northernmost and high elevation sampling sites over the northern parts of Canada, Europe, Siberia, and Mongolia. The database was dominated by areas where cool temperatures are not expected to be the main limiting factor on productivity, while the distribution of water limitation levels was more balanced.

We found a large bias towards conifers (81% of sites targeted conifer species). In fact, five conifers (*Pseudotsuga menziesii*, *Pinus sylvestris*, *Pinus ponderosa*, *Picea glauca*, and *Picea abies*) comprised almost a third of the ITRDB (29.6%) (Figure 4). This asymmetry also occurred within broadleaf trees. *Quercus* represented 69.3% of the broadleaf sites. Even in areas relatively well represented, such as the eastern US, most species were much less sampled than we would expect by their actual distribution (Figure S6). Less-common species, such as *Tsuga canadiensis*, *Pinus echinata*, or *Picea rubens*, dominated the database while data from some widespread species, such as *Quercus rubra* or *Pinus taeda*, were poorly represented.

ITRDB sites were commonly located in forests with low to medium vascular plant and forest diversity (Figure S4). Tropical areas were largely absent from the database while desert or semi-desert areas were also underrepresented.

3.3 | Priority sampling index

Priority assessments showed where the representativeness of the database can be improved with respect to several variables (Figure 5). In general, the global PSI index made it easier to interpret priorities (Figure 6). In the *Equal Weights* scenario, we found much of the entire African continent and tropical areas in South America, Asia, and Oceania to be a top priority. Filtering out areas with low seasonality and vegetation cover still showed large areas within those regions with high dendrochronological potential, such as dry tropical areas in central South America, central African savannas, and temperate or monsoon forests in south east Asia. *Ecological Research* priorities highlighted areas with high biodiversity in poorly sampled climatic regions while *Climatic Research* priorities also indicated the importance to update decades-old and widely distributed tree-ring records in Russia. (Figure 6).

4 | DISCUSSION

We took several steps to improve the usability of the ITRDB as a valuable resource for large-scale analyses on the climate and ecology of our planet. First, we provided a revised version of the database

to increase its accessibility and consistency for global meta-analyses. Second, we quantified important limitations and biases in the ITRDB, such as the alarming lack of information for Africa and low representation of tropical habitats, particularly in Asia, South America, and Oceania. We found that even well-represented regions were biased towards certain species and environments. Finally, to promote a more ecologically comprehensive ITRDB, we developed user-defined metrics to assist in the identification of priority areas for research.

4.1 | Current challenges in the use of the ITRDB

The current data format of ITRDB force researchers to manually correct errors (e.g., Breitenmoser et al., 2014) or omit large portions of data. St. George et al. (2013), for example, employed only 476 *Picea* sites of the 598 available. Importantly, corrections to the data are not always disclosed, limiting scientific reproducibility. Here, we tackled a first necessary step: creating a consistent, corrected, and updated version of the database, which we call the rITRDB. We hope the rITRDB will improve the accessibility and boost the use of dendrochronological data, foster discussions on dendro-data formats (e.g., Brewer, Murphy, & Jansma, 2011; Jansma, Brewer, & Zandhuis, 2010), and on big-data formatting and archiving in general.

The primary purpose for the ITRDB to date has been to just store dendrochronological data for climate reconstructions (Grissino-Mayer & Fritts, 1997). However, as dendroecological data are increasingly used in multidisciplinary studies, the shortcomings of the current dataset become more apparent, limiting the potential use of the database and, to a certain degree, the implications that can be drawn from it (Table 1). While some problems can be solved relatively easy within the current ITRDB framework (e.g., including extra metadata or changing data format, as shown here). Others stem from a deeper, more systemic root. Recent studies have stressed the importance of microtopography or genotype on tree growth, e.g., geomorphology (Bunn, Waggoner, & Graumlich, 2005; Lloyd, Sullivan, & Bunn, 2017; Piraino, Abraham, Diblasi, & Roig-Juñent, 2015), tree age and gender (Rozas, DeSoto, & Olano, 2009). A whole-community agreement is needed on how to incorporate this information to the ITRDB, including tree coordinates, individual tree characteristics, and other ring features, such as missing, false (inter-annual density fluctuations), and frost ring data. Sullivan and Csank (2016) suggested the incorporation of sampling design and other study characteristics in annexed files to the ITRDB metadata, which could alleviate this problem. In the long term, however, a deeper reconceptualization of the database is likely necessary, to include a wider range of data and metadata in the ITRDB.

Funding and publication pressure also play a crucial role in the choice of the species and locations targeted. We hope that the priority placed on open data by journals will increase data sharing in regions currently underrepresented in the ITRDB but that have active tree-ring research. Funding priorities are complex, and need not agree with the best choices for the community or the preferences of individual scientists. We hope that the PSI metric we develop here may prove useful for researches to present to funding agencies and scientific journals to justify the importance of increasing the diversity of species and environmental conditions studied with dendrochronological methods.

4.2 | Existing biases of the ITRDB data

4.2.1 | Spatial representativity

Our findings clearly suggest that the ITRDB is spatially imbalanced as approximately 75% of the database has been collected in North America and Europe, continents that constitute <25% of the world's land area. This bias can have a historical origin. Dendrochronological methods were developed in the early 20th Century in southwestern US and only branched out elsewhere by the mid to late 1900s. As the number of tree-ring scientists increased, data from different ecosystems have been added, resulting in a larger, though still patchy, coverage of global forest ecosystems. The bias towards European and North American habitats is a complex socioeconomical phenomenon that permeates all branches of ecology (Vellend et al., 2017). While it is accepted that most existing data syntheses represent an opportunistic collection of heterogeneous studies that do not necessarily represent natural ecosystems (e.g., Gonzalez et al., 2016; Yang, Ma, & Kreft, 2013), it is crucial that we start to quantify the biases and gaps in synthesis work to contextualize them and work towards more representative results.

A tree-ring database able to address global processes and mechanisms requires further information in African, eastern Asian, boreal Russian, Oceanian, and South American forests. This lack of data is aggravated by the recent reduction in submissions to the ITRDB (Babst et al., 2017). For this, it is crucial to ensure that tree-ring data continue being made public. For example, in all our scenarios, eastern Asia is highlighted as a high priority area. In fact, dendrochronological research has increased rapidly in the region since 2000, but only a small fraction of the data has been uploaded to the ITRDB (e.g., Shi, Li, Cook, Zhang, & Lu, 2012; Yang et al., 2014). This is certainly true for other regions. We encourage scientists and research organizations to make available this private data, which can greatly improve the representativity and diversity of the database overnight. A key aspect is the increasing push by scientific publishers and funders towards publicly accessible data. To take advantage of this to improve the ITRDB, we respectfully urge the ITRDB database to incorporate a DOI assignment for scientists to comply with journal requirements to use the ITRDB as data repository.

4.2.2 | Environmental representativity

The lack of tropical tree-ring data is the strongest bias regarding Earth's environmental variability. This is likely related to the long-standing assumption that low seasonality in tropical regions limits the formation of reliable tree-rings, an assumption that has been challenged over the last two decades (reviewed in Schöngart, Bräuning, Barbosa, Lisi, & de Oliveira, 2017). Tree-rings have become an emerging source of data to understand complex tropical ecosystems

(e.g., Paredes-Villanueva et al., 2013; Worbes, Herawati, & Martius, 2017). However, this has not translated yet into a better representation of tropical forests in the ITRDB. We hope that our more accessible version of the database also promotes the knowledge of the ITRDB database by tropical ecologists.

The seemingly low temperature sensitivity in the database may result in an overall underestimation of tree response to increasing temperatures by the ITRDB. This could be remedied by increasing the sampling of Arctic and sub-Arctic areas in North America and northern Eurasia (Churkina & Running, 1998).

The ITRDB represented mostly medium-to-low diversity forests, leaving the two extremes of the diversity gradient, tropical and arid regions, underrepresented. It could be important the lack of studies in the low end of the diversity gradient, as this is where endemic xerophilic species or populations are more likely to live. These species or populations are particularly sensitive to increasing temperatures and drought, and therefore face a significant threat from climatic change (Allen et al., 2010; van Mantgem et al., 2009) that is currently not captured by the ITRDB.

Condensing individual variable biases, defining priorities of research, and comparing importance between scenarios provided a useful common framework for the discussion and identification of strategic areas to sample to improve the ITRDB. We hope that our results encourage scientists to consider not only their research interests but also the representativity of the whole database when selecting their target areas and species. It is important to note, however, that PSI is not intended to define particular location studies, but rather priority areas. Topography, microhabitat, target species, local laws, administrative boundaries, accessibility, and human disturbance, among many others, need to be considered to define concrete sampling sites.

4.2.3 | Taxonomic representativity

From an ecological perspective, the taxonomic bias of the ITRDB is highly relevant. The database is dominated by conifers, and any conclusion from the database will mostly reflect coniferous ecology. For example, a recent meta-analysis using ITRDB data found an overall decrease in growth across North America (Charney et al., 2016). Eighty-five percent (85%) of their database, however, was comprised of coniferous site records. In diverse regions, conifers represent a small percentage of the total forest basal area (17.9% for trees with DBH >20 cm in eastern US, Figure S6). Taxonomic diversity needs to be considered to contextualize results and evaluate how they can be up-scaled to global mechanisms.

This taxonomic bias also extends into broadleaf species. *Quercus* dominate the database, but our comparison between forest inventory data and ITRDB sampling in eastern US similarly indicates that, while *Quercus* are important across the region, few species are sampled in the ITRDB and common broadleaf species are still highly underrepresented (Figure S6).

The classic sampling strategies in dendrochronology exacerbate the taxonomical bias, resulting in the underrepresentation of slow-growing and small trees and potentially leading to overestimations of forest productivity (discussed in Briffa & Melvin, 2011; Nehrbass-Ahles et al., 2014), spurious negative growth trends (Brienen et al., 2017), or divergence in climate–growth relationships (Alexander, 2017).

4.3 | Ways forward… a much-needed discussion

In an era of global change, big data, open access, and increasingly sophisticated computational tools, the field of dendrochronology and its primary database, the ITRDB, must evolve to meet the current scientific challenges. We have attempted here to both highlight the existing opportunities and provide resources to aid this endeavour. However, more steps are needed. An agreement on how to tackle the limitations of traditional tree-ring formats for storing metadata associated with these collections is urgently needed (see Babst et al., 2017; Breitenmoser et al., 2014; Brewer et al., 2011; Foster et al., 2014; Jansma et al., 2010). Also, the ITRDB needs to become a more comprehensive database, able to tackle complex relationships between tree growth and environment across scales. Formatting issues, up-to-date archiving services (such as DOI assignment to data contributions), and increased economic support for data curators, will ensure the quality of newly incorporated data and the improvement of the ITRDB.

Adding information on the sampling design on the current ITRDB (Sullivan & Csank, 2016) would improve the dataset for ecological syntheses, especially if these data are merged into the files, rather than annexed to them. Of similar importance is integrating the ITRDB data with other existing datasets with tree-ring information (e.g., TRY or Forest Inventory and Analysis (FIA) National Program), to centralize all tree-ring data. This, however, requires multi-institutional coordination, including governmental agencies. A first step in this direction could be to show the FIA plot locations as ghost points in the current ITRDB, for users to know the information available in parallel repositories that they can include in their research.

Increasing the diversity of sampling locations and species requires effort and funding. Convincing funding or publication agencies of the importance of a representative ITRDB could be a quixotic task. PSI can help to stress the importance or novelty of underrepresented areas. Other low-cost activities, however, could also contribute to this end. First, we suggest that dendrochronological workshops and fieldweeks could be used to target underrepresented areas and species. Products from these efforts could then be made publicly available. Second, scientific publishers and funders (particularly those related to the Tree Ring Society) could demand publicly accessible data in their associated publications. We stress here again the importance of incorporating DOI assignment in the ITRDB to achieve this goal, as currently submission to ITRDB does not guarantee compliance with the open data policies of many scientific journals. Also, results from forests less sensitive to environmental conditions may contain complex signals and more statistical noise, making their publication harder. We encourage reviewers and editors to consider the value of this information to improve our understanding of the ecology, and climate change responses of natural forests. Finally, we are aware of large collections of existing data outside the

ITRDB in relatively obscure literature, PhD and MSc theses, or "collecting dust" in private computers. We could not encourage researchers more to share these valuable data for the benefit of the whole scientific community, and dendrochronology in particular.

Dendrochronological data have already shown its potential to address long-standing ecological questions in innovative ways, such as the diversity-stability debate (Jucker et al., 2014), the influence of environmental conditions in the relationship between biodiversity and ecosystem functions (Grossiord, Granier, Gessler, Jucker, & Bonal, 2014), or the intricate genetic factors affecting adaptation to local conditions (Housset et al., 2018). But it is not fortuity that they used detailed individual and plot level data rather than the ITRDB. More, and more diverse, tree-ring data are needed to create a more representative ITRDB that will provide a global understanding of climate change, climate science, and ecological processes.

ACKNOWLEDGEMENTS

DATA ACCESSIBILITY

As stated in the main text, the reformatted ITRDB dataset is available via NOAA web repository https://www.ncdc.noaa.gov/paleo/study/25570. All topographic and environmental layers are publicly available in the links provided in Table 2. The R script and data used for our analyses are available via dryad repository https://doi.org/10.5061/dryad.kh0qh06.

Title: Data from: The International Tree-Ring Data Bank (ITRDB) revisited: data availability and global ecological representativity

DOI: https://doi.org/10.5061/dryad.kh0qh06

Journal: Journal of Biogeography

Journal manuscript number: JBI-18-0195

ORCID

Rubén D. Manzanedo (iD) http://orcid.org/0000-0001-6592-7235

REFERENCES

Alexander, M. R. (2017). *Determining the role of stand structure in shaping climate-growth relationships in eastern temperate forests of the US*. Tucson, AZ: The University of Arizona.

Allen, C. D., Macalady, A. K., Chenchouni, H., Bachelet, D., McDowell, N., Vennetier, M., & Cobb, N. (2010). A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *Forest Ecology and Management*, *259*, 660–684. https://doi.org/10.1016/j.foreco.2009.09.001

Amoroso, M. M., Daniels, L. D., Baker, P. J., & Camarero, J. J. (2017). Introduction. In M. M. Amoroso, L. D. Daniels, P. J. Baker, & J. J. Camarero (Eds.), *Dendroecology: Tree-ring analyses applied to ecological studies*. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-61669-8

Babst, F., Poulter, B., Bodesheim, P., Mahecha, M. D., & Frank, D. C. (2017). Improved tree-ring archives will support earth-system science. *Nature Ecology and Evolution*, *1*, 0008.

Blume-Werry, G., Kreyling, J., Laudon, H., & Milbau, A. (2016). Short-term climate change manipulation effects do not scale up to long-term legacies: Effects of an absent snow cover on boreal forest plants. *Journal of Ecology*, *104*, 1638–1648. https://doi.org/10.1111/1365-2745.12636

Breitenmoser, P., Brönnimann, S., & Frank, D. (2014). Forward modelling of tree-ring width and comparison with a global network of tree-ring chronologies. *Climate of the Past*, *10*, 437–449. https://doi.org/10.5194/cp-10-437-2014

Brewer, P. W., Murphy, D., & Jansma, E. (2011). Tricycle: A universal conversion tool for digital tree-ring data. *Tree-Ring Research*, *67*, 135–144. https://doi.org/10.3959/2010-12.1

Brienen, R. J. W., Gloor, M., & Ziv, G. (2017). Tree demography dominates long-term growth trends inferred from tree rings. *Global Change Biology*, *23*, 474–484. https://doi.org/10.1111/gcb.13410

Briffa, K. R., & Melvin, T. M. (2011). A closer look at regional curve standardisation of tree-ring records: Justification of the need, a warning of some pitfalls, and suggested improvements in its application. In M. K. Hughes, T. W. Swetnam, & H. F. Diaz (Eds.), *Dendroclimatology: Progress and prospects* (pp. 113–145). Dordrecht, Netherlands: Springer. https://doi.org/10.1007/978-1-4020-5725-0

Bunn, A. G. (2008). A dendrochronology program library in R (dplR). *Dendrochronologia*, *26*, 115–124. https://doi.org/10.1016/j.dendro.2008.01.002

Bunn, A. G., Waggoner, L. A., & Graumlich, L. J. (2005). Topographic mediation of growth in high elevation foxtail pine (*Pinus balfouriana* Grev. et Balf.) forests in the Sierra Nevada, USA. *Global Ecology and Biogeography*, *14*, 103–114. https://doi.org/10.1111/j.1466-822X.2005.00141.x

Charney, N. D., Babst, F., Poulter, B., Record, S., Trouet, V. M., Frank, D., & Evans, M. E. (2016). Observed forest sensitivity to climate implies large changes in 21st century North American forest growth. *Ecology letters*, *19*, 1119–1128. https://doi.org/10.1111/ele.12650

Churkina, G., & Running, S. W. (1998). Contrasting climatic controls on the estimated productivity of global terrestrial biomes. *Ecosystems*, *1*, 206–215. https://doi.org/10.1007/s100219900016

Cook, E. R., Woodhouse, C. A., Eakin, C. M., Meko, D. M., & Stahle, D. W. (2004). Long-term aridity changes in the western United States. *Science*, *306*, 1015–1018. https://doi.org/10.1126/science.1102586

Core Team, R. (2017). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Davis, S. C., Hessl, A. E., Scott, C. J., Adams, M. B., & Thomas, R. B. (2009). Forest carbon sequestration changes in response to timber harvest. *Forest Ecology and Management*, *258*, 2101–2109. https://doi.org/10.1016/j.foreco.2009.08.009

Foster, J. R., D'Amato, A. W., & Bradford, J. B. (2014). Looking for age-related growth decline in natural forests: Unexpected biomass patterns from tree rings and simulated mortality. *Oecologia*, *175*, 363–374. https://doi.org/10.1007/s00442-014-2881-2

Fritts, H. C., Blasing, T. J., Hayden, B. P., & Kutzbach, J. E. (1971). Multivariate techniques for specifying tree-growth and climate relationships and for reconstructing anomalies in paleoclimate. *Journal of*

Applied Meteorology, 10, 845–864. https://doi.org/10.1175/1520-0450(1971)010<0845:MTFSTG>2.0.CO;2

Fritts, H. C., & Swetnam, T. W. (1989). Dendroecology: A tool for evaluating variations in past and present forest environments. *Advances in Ecological Research*, 19, 111–188. https://doi.org/10.1016/S0065-2504(08)60158-0

Gedalof, Z.E., & Berg, A.A. (2010). Tree ring evidence for limited direct CO_2 fertilization of forests over the 20th century. *Global Biogeochemical Cycles*, 24, GB3027.

Gonzalez, A., Cardinale, B. J., Allington, G. R., Byrnes, J., Arthur Endsley, K., Brown, D. G., & Loreau, M. (2016). Estimating local biodiversity change: A critique of papers claiming no net loss of local diversity. *Ecology*, 97, 1949–1960. https://doi.org/10.1890/15-1759.1

Graumlich, L. J., & Brubaker, L. B. (1986). Reconstruction of annual temperature (1590–1979) for Longmire, Washington, derived from tree rings. *Quaternary Research*, 25, 223–234. https://doi.org/10.1016/0033-5894(86)90059-1

Grissino-Mayer, H. D., & Fritts, H. C. (1997). The International Tree-Ring Data Bank: An enhanced global database serving the global scientific community. *The Holocene*, 7, 235–238. https://doi.org/10.1177/095968369700700212

Grossiord, C., Granier, A., Gessler, A., Jucker, T., & Bonal, D. (2014). Does Drought Influence the Relationship Between Biodiversity and Ecosystem Functioning in Boreal Forests? *Ecosystems*, 17, 394–404. https://doi.org/10.1007/s10021-013-9729-1

Heinrich, I., & Allen, K. (2013). Current issues and recent advances in Australian dendrochronology: Where to next? *Geographical Research*, 51, 180–191. https://doi.org/10.1111/j.1745-5871.2012.00786.x

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25, 1965–1978. https://doi.org/10.1002/(ISSN)1097-0088

Housset, J. M., Nadeau, S., Isabel, N., Depardieu, C., Duchesne, I., Lenz, P., & Girardin, M. P. (2018). Tree rings provide a new class of phenotypes for genetic associations that foster insights into adaptation of conifers to climate change. *New Phytologist*, 218, 630–645. https://doi.org/10.1111/nph.14968

Jacoby, G. C., & D'Arrigo, R. (1989). Reconstructed Northern Hemisphere annual temperature since 1671 based on high-latitude tree-ring data from North America. *Climatic Change*, 14, 39–59. https://doi.org/10.1007/BF00140174

Jansma, E., Brewer, P. W., & Zandhuis, I. (2010). TRiDaS 1.1: The tree-ring data standard. *Dendrochronologia*, 28, 99–130. https://doi.org/10.1016/j.dendro.2009.06.009

Jucker, T., Bouriaud, O., Avacaritei, D., & Coomes, D. A. (2014). Stabilizing effects of diversity on aboveground wood production in forest ecosystems: Linking patterns and processes. *Ecology Letters*, 17, 1560–1569. https://doi.org/10.1111/ele.12382

Kier, G., Mutke, J., Dinerstein, E., Ricketts, T. H., Küper, W., Kreft, H., & Barthlott, W. (2005). Global patterns of plant diversity and floristic knowledge. *Journal of Biogeography*, 32, 1107–1116. https://doi.org/10.1111/j.1365-2699.2005.01272.x

Kottek, M., Grieser, J., Beck, C., Rudolf, B., & Rubel, F. (2006). World Map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift*, 15, 259–263. https://doi.org/10.1127/0941-2948/2006/0130

Krakauer, N.Y., & Randerson, J.T. (2003). Do volcanic eruptions enhance or diminish net primary production? *Evidence from tree rings. Global Biogeochemical Cycles*, 17(4).

Lloyd, A. H., Sullivan, P. F., & Bunn, A. G. (2017). Integrating dendroecology with other disciplines improves understanding of upper and latitudinal treelines. In M. M. Amoroso, L. D. Daniels, P. J. Baker, & J. J. Camarero (Eds.), *Dendroecology: Tree-ring analyses applied to ecological studies*. Cham: Springer International Publishing.

Mina, M., Martin-Benito, D., Bugmann, H., & Cailleret, M. (2016). Forward modeling of tree-ring width improves simulation of forest growth responses to drought. *Agricultural and Forest Meteorology*, 221, 13–33. https://doi.org/10.1016/j.agrformet.2016.02.005

Myers-Smith, I. H., Forbes, B. C., Wilmking, M., Hallinger, M., Lantz, T., Blok, D., & Lévesque, E. (2011). Shrub expansion in tundra ecosystems: Dynamics, impacts and research priorities. *Environmental Research Letters*, 6, 045509. https://doi.org/10.1088/1748-9326/6/4/045509

Nehrbass-Ahles, C., Babst, F., Klesse, S., Nötzli, M., Bouriaud, O., Neukom, R., & Frank, D. (2014). The influence of sampling design on tree-ring-based quantification of forest growth. *Global Change Biology*, 20, 2867–2885. https://doi.org/10.1111/gcb.12599

O'Donnel, M.S., & Ignizio, D.A. (2012). Bioclimatic predictors for supporting ecological applications in the conterminous United States: U.S. Geological Survey Data Series 691.

Paredes-Villanueva, K., Sánchez-Salguero, R., Manzanedo, R. D., Sopepi, R. Q., Palacios, G., & Navarro-Cerrillo, R. M. (2013). Growth rate and climatic response of Machaerium scleroxylon in a dry tropical forest in southeastern Santa Cruz, Bolivia. *Tree-Ring Research*, 69, 63–79. https://doi.org/10.3959/1536-1098-69.2.63

Pederson, N., Dyer, J. M., McEwan, R. W., Hessl, A. E., Mock, C. J., Orwig, D. A., & Cook, B. I. (2014). The legacy of episodic climatic events in shaping temperate, broadleaf forests. *Ecological Monographs*, 84, 599–620. https://doi.org/10.1890/13-1025.1

Piraino, S., Abraham, E. M., Diblasi, A., & Roig-Juñent, F. A. (2015). Geomorphological-related heterogeneity as reflected in tree growth and its relationships with climate of Monte Desert *Prosopis flexuosa* DC woodlands. *Trees*, 29, 903–916. https://doi.org/10.1007/s00468-015-1173-8

Rozas, V., DeSoto, L., & Olano, J. M. (2009). Sex-specific age-dependent sensitivity of tree-ring growth to climate in the dioecious tree *Juniperus thurifera*. *New Phytologist*, 182, 687–697. https://doi.org/10.1111/j.1469-8137.2009.02770.x

Schmidt, N. M., Baittinger, C., Kollmann, J., & Forchhammer, M. C. (2010). Consistent Dendrochronological Response of the Dioecious Salix arctica to Variation in Local Snow Precipitation across Gender and Vegetation Types. *Arctic, Antarctic, and Alpine Research*, 42, 471–475. https://doi.org/10.1657/1938-4246-42.4.471

Schöngart, J., Bräuning, A., Barbosa, A. C. M. C., Lisi, C. S., & de Oliveira, J. M. (2017). Dendroecological studies in the neotropics: History, status and future challenges. In M. M. Amoroso, L. D. Daniels, P. J. Baker, & J. J. Camarero (Eds.), *Dendroecology: Tree-ring analyses applied to ecological studies*. Cham: Springer International Publishing.

Schweingruber, F. H., Bräker, O. U., & Schär, E. (1979). Dendroclimatic studies on conifers from central Europe and Great Britain. *Boreas*, 8, 427–452.

Shi, J., Li, J., Cook, E. R., Zhang, X., & Lu, H. (2012). Growth response of *Pinus tabulaeformis* to climate along an elevation gradient in the eastern Qinling Mountains, central China. *Climate Research*, 53, 157–167. https://doi.org/10.3354/cr01098

Simard, M., Pinto, N., Fisher, J. B., & Baccini, A. (2011). Mapping forest canopy height globally with spaceborne lidar. *Journal of Geophysical Research: Biogeosciences*, 116, G04021.

St. George, S., Ault, T.R., & Torbenson, M.C.A. (2013). The rarity of absent growth rings in Northern Hemisphere forests outside the American Southwest. *Geophysical Research Letters*, 40, 3727–3731. https://doi.org/10.1002/grl.50743

Stahle, D. W., Cleaveland, M. K., & Hehr, J. G. (1985). A 450-year drought reconstruction for Arkansas, United States. *Nature*, 316, 530–532. https://doi.org/10.1038/316530a0

Stockton, C., & Jacoby, G. (1976). Long-term surface water supply and streamflow trends in the Upper Colorado River Basin. In *Lake Powell Research Project Bulletin No. 18*. Arlington, VA: Nation Science Foundation.

Stoffel, M., & Bollschweiler, M. (2008). Tree-ring analysis in natural hazards research? an overview. *Natural Hazards and Earth System Science*, *8*, 187–202. https://doi.org/10.5194/nhess-8-187-2008

Sullivan, P. F., & Csank, A. Z. (2016). Contrasting sampling designs among archived datasets: Implications for synthesis efforts. *Tree physiology*, *36*, 1057–1059. https://doi.org/10.1093/treephys/tpw067

Tei, S., Sugimoto, A., Yonenobu, H., Matsuura, Y., Osawa, A., Sato, H., & Maximov, T. (2017). Tree-ring analysis and modeling approaches yield contrary response of circumboreal forest productivity to climate change. *Global Change Biology*, *23*, 5179–5188. https://doi.org/10.1111/gcb.13780

Van Mantgem, P. J., Stephenson, N. L., Byrne, J. C., Daniels, L. D., Franklin, J. F., Fulé, P. Z., & Veblen, T. T. (2009). Widespread Increase of Tree Mortality Rates in the Western United States. *Science*, *323*, 521–524. https://doi.org/10.1126/science.1165000

Vellend, M., Dornelas, M., Baeten, L., Beauséjour, R., Brown, C. D., De Frenne, P., & Myers-Smith, I. H. (2017). Estimates of local biodiversity change over time stand up to scrutiny. *Ecology*, *98*, 583–590. https://doi.org/10.1002/ecy.1660

Vicente-Serrano, S. M., Camarero, J. J., & Azorin-Molina, C. (2014). Diverse responses of forest growth to drought time-scales in the Northern Hemisphere. *Global Ecology and Biogeography*, *23*, 1019–1030. https://doi.org/10.1111/geb.12183

Wiens, J. A. (1989). Spatial Scaling in Ecology. *Functional Ecology*, *3*, 385–397. https://doi.org/10.2307/2389612

Williams, A. P., Allen, C. D., Macalady, A. K., Griffin, D., Woodhouse, C. A., Meko, D. M., & McDowell, N. G. (2013). Temperature as a potent driver of regional forest drought stress and tree mortality. *Nature Climate Change*, *3*, 292–297. https://doi.org/10.1038/nclimate1693

Williams, A. P., Allen, C. D., Millar, C. I., Swetnam, T. W., Michaelsen, J., Still, C. J., & Leavitt, S. W. (2010). Forest responses to increasing aridity and warmth in the southwestern United States. *Proceedings of the National Academy of Sciences*, *107*, 21289–21294. https://doi.org/10.1073/pnas.0914211107

Wilson, R., D'arrigo, R., Buckley, B., Büntgen, U., Esper, J., Frank, D., … Youngblut, D. (2007). A matter of divergence: Tracking recent warming at hemispheric scales using tree ring data. *Journal of Geophysical Research*, *117*(D17).

Worbes, M., Herawati, H., & Martius, C. (2017). Tree Growth Rings in Tropical Peat Swamp Forests of Kalimantan. *Indonesia. Forests*, *8*, 336. https://doi.org/10.3390/f8090336

Yang, W., Ma, K., & Kreft, H. (2013). Geographical sampling bias in a large distributional database and its effects on species richness–environment models. *Journal of Biogeography*, *40*, 1415–1426. https://doi.org/10.1111/jbi.12108

Yang, B., Qin, C., Wang, J., He, M., Melvin, T. M., Osborn, T. J., & Briffa, K. R. (2014). A 3,500-year tree-ring record of annual precipitation on the northeastern Tibetan Plateau. *Proceedings of the National Academy of Sciences*, *111*, 2903–2908. https://doi.org/10.1073/pnas.1319238111

Zang, C. (2015). Dendrobox – An interactive exploration tool for the International Tree Ring Data Bank. *Dendrochronologia*, *33*, 31–33. https://doi.org/10.1016/j.dendro.2014.10.002

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.