Robust Historical Forest Inventories Describe Patterns of Forest Structure, Composition, and Competition across a large latitudinal gradient

Connie Ryan*, Brandon Collins, R. Keala Hagmann, Robert York

Affiliations

*Corresponding author email: cpryan@berkeley.edu

1. Abstract

The structure and composition of historical frequent fire forests in California and Southern Oregon has been inferred from a variety of sources, including systematic landscape-scale timber inventories that provide robust estimates of historical forest conditions. We compiled for the first time five such inventories, collected between 1911-1936 and collectively representing over 20,000 individual belt transects, to summarize historical forest structure and composition of ponderosa pine and mixed-conifer forests across the Sierra Nevada of California and the Southern Cascades of Oregon. Furthermore, we assessed the effects of common data filtering practices, including removal of areas of low stocking and of small-diameter trees, on resulting structure and composition estimates; we additionally compared different approaches for calculating a metric of inter-tree competition, the stand density index (SDI). Finally, we characterized broad patterns of forest structure and composition across the wide latitudinal gradient represented in the data by cluster identification through K-means clustering. These historical forests, on average, contained low tree densities and stocking, moderate to large tree sizes, and low levels of competition, yet we identified substantial variability across the landscape likely corresponding to variation in local environmental conditions and fire regimes. Areas of low stocking varied in frequency across the landscape, and their removal increased estimates of tree density, stocking, and competition, but showed minimal effect on estimates of tree size. Small trees made up substantial portions of historical tree density, and their removal showed little effect on estimates of stocking and competition, but increased estimates of tree size and decreased estimates of tree density. Calculation of SDI through the 'traditional' method was shown to slightly overestimate competition by approximately 2-4%. Finally, clustering analysis identified high pine fraction as a unique characteristic of forests in the Southern Cascades of Oregon, though shared structural and compositional types were found across the latitudinal gradient represented in the data, likely due to the divergent influence on forest structure and composition of gradients of site productivity, in tandem with the homogenizing influence of frequent fire. These results build on evidence suggesting historical forests were characterized by distinctly lower levels of competition than contemporary forests, and characterize the wide range of conditions seen across historical forested landscapes that can be used to guide forest management in creation of resilient forests with a diversity of stand structures and compositions tailored to local site conditions.

Keywords: Quarter-Quarter, timber inventory, stand density index, forest structure and composition, frequent fire, resilience

2. Introduction

Forests in western North America have undergone substantial shifts in structure and composition since the advent of Euro-American colonization in the mid-19th century, primarily through the suppression of Indigenous and lightning-ignited fire as well as substantial logging and land development (Taylor et al. 2016, Hagmann et al. 2021). While climate change poses novel challenges to these forests, in addition to those caused by interruption of fire regimes, information about the processes structuring pre-colonization (historical) forests in western North America can inform management efforts to preserve forest ecosystem health and ultimately prevent wide-scale loss of forested areas (Safford et al. 2012). Mimicking the effects of key ecosystem processes from the historical period, such as frequent fire that reduced tree densities and maintained high tree vigor, shows promise in guiding management interventions to produce ecosystem characteristics, such as resilience to stressors, that were common in the historical period and are desirable in contemporary forests (North et al. 2022). Information about historical processes can be inferred from a variety of sources, but detailed quantitative data describing historical forest structure and composition frequently provide a starting point from which the characteristics of historical processes can be inferred (Hagmann et al. 2021).

While a variety of data sources exist that describe historical forest structure and composition, including reconstructions of past forest conditions from contemporary old growth trees and stumps (Taylor 2004, North et al. 2007, Scholl and Taylor 2010, Van de Water and North 2011, Sensenig et al. 2013, Maxwell et al. 2014, Taylor et al. 2014, Barth et al. 2015), small-scale (<10 hectare (ha)) forest inventories (Knapp et al. 2013), landscape-scale PLSS witness tree records (Knight et al. 2020), and landscape-scale plot networks (Wieslander 1935, McIntyre et al. 2015), each have limitations. Reconstructions and small-scale inventories generally represent areas too small to assess landscape-scale patterns of structure and composition, and even intensive historical inventories that thoroughly census larger areas (e.g. ~4000 ha, Ritchie (2016)) do not capture the landscape-scale. PLSS witness tree records span large areas, but represent extremely low sampling intensity, while large-scale plot networks have non-random locations and often have sampling designs that prevent the accurate characterization of forest structure (Wieslander 1935, Knight et al. 2020).

This paper synthesizes data from five early 20th century (1911-1936) timber inventories conducted across the Sierra Nevada in California and in the southern Cascade Range in Oregon by government agencies and private companies to assess stocking of timber through transects systematically arranged across the Bureau of Land Management's Public Land Survey System (PLSS) (Hagmann et al. 2013, Collins et al. 2015, 2021, Stephens et al. 2015, 2018, 2023). These data represent the most robust form of quantitative information describing the structure and composition of these forests in the early 20th century, and while the majority of these data have been lost, they do not suffer from the same limitations described for other sources. They report data at the landscape-scale, have systematic sampling, and represent reasonably intensive sampling efforts that capture variability across the landscape (Hagmann et al. 2018). However, these data do contain their own limitations, centering on exclusion of smaller trees, non-conifer species, and dead trees, all of which were not of interest for timber production. These data have been described individually and occasionally in conjunction (North et al. 2022) but have not been summarized together. This presented an opportunity to build on existing work describing historical structure and composition across a wide latitudinal gradient of frequent fire forests in California and Southern Oregon and explore whether frequent fire maintained similar forest structure and composition across this wide gradient of soil type/parent material, precipitation, temperature, and topographical relief.

Additionally, these data allow an assessment of the competitive environment experienced by these forests under an undisturbed fire regime through the calculation of relative density metrics (Chivhenge et al. 2024). Initial evidence summarized from a portion of the data presented here suggests that historical frequent fire forests experienced drastically lower levels of competition, characterized by the common forestry metric Stand Density Index (SDI) than their contemporary counterparts (North et al. 2022). The implied connection between low levels of competition and the resilience of frequent fire forests over millennia has significant implications for forest management and lends utility to sources of data characterizing SDI in historical frequent fire forests. While a variety of sources can be used to do this, systematic inventories such as those presented here can robustly characterize variability in SDI, and therefore competitive dynamics, at moderate to fine resolution across the landscape-scale. However, SDI can be calculated through several similar equations that produce differing results, with the summation or additive method known to produce more accurate values in stands with non-normal diameter distributions when compared to the traditional method (Long and Daniel 1990). For numerous sources of historical data, including several datasets described here, formatting of data prevents calculation of additive SDI, which is known to be bounded between approximately 80-100% of traditional SDI values when calculated for the same stand (Ducey 2009). As the ratio of additive to traditional SDI for a given stand is known to be a property of the shape of the stand's diameter distribution, the ability to calculate additive SDI for several datasets described here provides an opportunity to characterize the likely overestimation of SDI when calculated with the traditional method for the characteristic diameter distributions of these historical forests.

Finally, individual studies describing these datasets have applied a variety of filters to the data before analysis, including removal of areas with low basal area (Stephens et al. 2018, Collins et al. 2021) and varying minimum diameter thresholds. The large amount of data here allowed for assessment of the effects of these analysis choices on resulting structure and composition metrics, to better characterize the likely effects of these choices on descriptions of historical forests.

Altogether, access to the data presented here represented an opportunity to describe historical structure, composition, and competitive environment of frequent fire forests across a wide latitudinal gradient, as well as to characterize the potential error introduced by method of SDI calculation and the effects of common analysis choices on structure and composition metrics.

3. Methods

3.1 Study Site

Data were collected across private, public, and reservation (publicly held trust) lands in the Sierra Nevada mountains of California and in the Cascade Range in south-central Oregon, and approximately span latitudes 35.6ºN-43.4ºN and elevations 1058m-2444m. Forest types represented in these data capture the majority of frequent fire montane forests in California and Southern Oregon, including ponderosa pine (*Pinus ponderosa*) and mixed conifer forests. Climate across sites is generally Mediterranean, with hot dry summers and cool wet winters with variable proportions of precipitation as snow. Historically, fire was frequent in these forests, with ignitions from both Indigenous peoples and lightning strongly influencing forest structure and composition (Anderson 2005). Across datasets, elevation, soils, and precipitation vary, while fire return interval (FRI) is more stable.

The southernmost dataset was collected in the Greenhorn Mountains of the southern Sierra Nevada on the former Kern National Forest (now Sequoia National Forest) (KERN) (Stephens et al. 2015). Forests in KERN are characterized as ponderosa pine and mixed-conifer forests, with species including ponderosa pine, sugar pine (*P. lambertiana*), white fir *(Abies concolor*), incense-cedar (*Calocedrus decurrens*), California black oak (*Quercus kelloggii*), and canyon live oak (*Q. chrysolepis*). Annual precipitation is lower here than other study sites, with a mean 32 cm yr^{-1}. Temperatures range from mean values of 0.9ºC in January to 20.3ºC in July (data from Piutes Remote Automated Weather Station, 2005-2011). Elevation is high relative to other study sites, with a mean of 1833m and range 1467m-2211m. This area experienced historically frequent fire, with a mean point return interval of ~5-20 years (Kilgore and Taylor 1979, Caprio and Swetnam 1995).

Data from the central Sierra Nevada were collected on the Stanislaus National Forest and a portion of Yosemite National Park (STAN) (Collins et al. 2015). Forests in STAN are characterized as pine-mixed-conifer, consisting of sugar pine, ponderosa pine, white fir, red fir (*A. magnifica*), incense-cedar, and Douglas-fir *(Pseudotsuga menziesii*). Annual precipitation is moderate relative to other datasets, falling primarily as snow and averaging just above 100cm yr^{-1}. Temperatures range from mean values of 4ºC in January to 20ºC in July (data from Crane Flat Remote Automated Weather Station, 1999-2014). Elevation in low compared to other datasets, with a mean of 1432m and range 1058m-2147m. This area experienced historically frequent fire, with a mean point return interval of 12 years (Scholl and Taylor 2010).

Data from the northern Sierra Nevada were collected on the El Dorado National Forest (ELDO) (Stephens et al. 2018). Forests in ELDO are characterized as fir-dominated mixed-conifer, consisting of white fir, incense-cedar, sugar pine, Jeffrey pine (*P. jeffreyi*), and ponderosa pine, with small amounts of California black oak, lodgepole pine (*P. contorta*), and Douglas-fir, and occasional western juniper (*Juniperus occidentalis*), western white pine (*P. monticola*), and canyon live oak. Precipitation here is highest of study sites, averaging 140 cm yr^{-1}. Temperatures average 15.4ºC annually (Flint et al. 2013). Elevation is high relative to other datasets, with a mean of 1846m and range 1199m-2444m. Fire regimes in this area likely included a range of fire return intervals; median composite interval at the 9-15ha scale was 4.7 years in a nearby pine-mixed-conifer forest, but as this forest represents a transition from mixed-conifer to true fir forest, moderate-length intervals of 20 to 30 years were also likely common (Stephens and Collins 2004, Meyer and North 2019).

Data from the southern Cascade Range and a portion of the northern Sierra Nevada were collected on land owned now by Collins Pine Company (COLLINS) (Collins et al. 2021). Forests in COLLINS are characterized as mixed-conifer, consisting of ponderosa pine, sugar pine, incense-cedar and Douglas-fir, with small amounts of red fir. Precipitation is high, averaging 124 cm yr^{-1}. Temperatures range from mean values of -3.9ºC in January to 27.3ºC in July, with annual mean of 15.4ºC (Flint et al. 2013). Elevation is moderate compared to other datasets, with mean 1554m and range 1102m-2102m. Fire was historically frequent in this area, with median fire return intervals of 12-14 years (Skinner et al. 2018).

Data from the Cascade Range in south-central Oregon were collected on the former Klamath Indian Reservation (now Fremont-Winema National Forest) (KIR) (Hagmann et al. 2013). Forests in KIR are characterized as ponderosa pine and mixed-conifer forests, with small areas of lodgepole pine and upper montane forests. Both forest types in this site are dominated by ponderosa pine, though ponderosa pine forests occasionally contained small amounts of lodgepole pine, and mixed conifer forest contained ponderosa pine, white fir, lodgepole pine, and sugar pine, with smaller amounts of Douglas-fir, incense-cedar, western juniper, western white pine, noble fir (*A. procera*), mountain hemlock (*Tsuga mertensiana*), and whitebark pine (*P. albicaulis*). Precipitation is relatively low compared to other sites, with mean annual precipitation around 60-70cm yr^{-1} (Hagmann et al. 2019, Appendix B). Temperatures range from means of -6ºC to -8ºC in January to 25ºC to 28ºC in July (PRISM, 2006). Elevation is low compared to other datasets, averaging 1400m but ranging 1270-2300m. Fire was historically frequent and low-severity, with a fire return interval of 14.9 years, although stand-replacing fires were common in lodgepole pine forests and upper montane forests likely experienced less frequent and higher severity fires (Hagmann et al. 2019).

3.2 Data Collection

Original cruises conducted by the USDA United States Forest Service (USFS) took place in 1911 (STAN, KERN) and 1923-36 (ELDO). Cruises conducted by Brown and Brown Inc. for Curtis, Collins & Holbrook Company on land now owned by Collins Pine Company took place in 1924 (COLLINS), and cruises conducted by the US Indian Service (now the Bureau of Indian Affairs) took place 1914-1922 (KIR). Data were accessed by collaborators primarily from the National Archives and Records Administration, but data collected on private land (COLLINS) was provided by the current owners (Collins Pine Company) (NARA, 1914-1922). Scans of historical cruise datasheets were manually digitized by collaborators.

Data collected in these historical inventories centered on trees, including species and diameter at breast height (DBH), with diameter recorded to the nearest 5.1 cm (2 inches). Other data were often recorded but are not analyzed here, including tree height, shrub cover, and written notes describing a variety of site conditions. Data were collected in belt transects, with transect locations tied to PLSS quarter-quarter sections (16.2 ha, 40 acre (ac)), ensuring systematic sampling across the landscape. Due to inherent variation in landscape conditions (e.g. water bodies), sections were occasionally divided into 'lots' of variable shape and size in place of square quarter-quarter sections, but lots are referred to here as quarter-quarter sections.

Transects were generally 402 meters (m) in length (20 chains (ch) or 1320 feet (ft), the length of a quarter-quarter section) and varied from 20.1m (1 ch, 66 ft) to 40.2m (2 ch, 132 ft) in width (Table 1). Each quarter-quarter section contained between 1 and 4 transects, producing sampling intensities of 10-40%. Cruises recorded all live conifers above either 15.2 centimeters (cm) or 30.5 cm (6 inches (in) or 12 in), with inconsistent recording of non-conifers and dead trees between inventories. Quality control assessments ("check cruises") conducted at the time of original surveys on a subset of surveyed area showed high accuracy of original cruises, with slight overestimation (4-11%) of tree densities in original surveys on average (Hagmann et al. 2018).

Due to variation in original cruise layouts and how original datasheets were accessed and digitized, the resolution and spatial scale of the data varies. STAN, KERN, and KIR report tree-level data, while ELDO and COLLINS report averages of tree-level records. STAN, KERN, ELDO, and COLLINS all report data at the quarter-quarter scale, with data for each quarter-quarter collected in 1-4 transects (Table 1). KIR reports data at scales varying from the half-quarter-quarter (8.1ha, 20ac) to the quarter (64.7ha, 160ac), with data for each sampling unit collected in 1-4 transects. To simplify reference, the smallest spatial unit for which data is reported is referred to as the "sample"; for STAN, KERN, ELDO, and COLLINS this is the quarter-quarter, while for KIR this varies from the half-quarter-quarter to the quarter. The sampling methods described above characterize the vast majority of data in these inventories, but small amounts of variation from these procedures did occur. For complete sampling details, see: for KERN, Stephens et al. (2015); for STAN, Collins et al. (2015); for ELDO, Stephens et al. (2018); for COLLINS, Collins et al. (2021); for KIR, Hagmann et al. (2013).

3.3 Data Preparation

All data preparation and cleaning, analysis, and visualization took place in RStudio (v. 2023.12.1+402) running R (v. 4.3.3), utilizing "tidyverse" packages, package "patchwork", and package "viridis" (Wickham et al. 2019, Posit Team 2022, R Core Team 2022, Garnier et al. 2024, Pedersen 2024).

Inventory data were received from collaborators already digitized as spreadsheets, frequently with data preparation steps already completed. See original papers referenced in Sections 3.1 and 3.2 for complete details; data preparation steps described here were not completed in data shared by collaborators. Formatting of data received from collaborators varied between datasets. Data in COLLINS consisted of tree density and mean diameter estimates by species for each sample; data in ELDO consisted of tree density estimates in 15.2cm classes, and basal area estimates by species for each sample; data in KERN and STAN consisted of individual tree records with diameter and species information for each sample; and data in KIR consisted of tree-level records with unique diameter and species information that had been summarized from individual-tree records, such that each record represented varying numbers of individual trees, for each sample. All datasets also reported information describing transect sizes in which data were collected, allowing calculation of per-area values.

For all datasets, dead trees and non-conifers were either not measured originally or removed by collaborators; however, due to formatting of original datasheets for ELDO, removal of non-conifers was not possible and they are included in this analysis, but represent a minor component (<2% of total basal area). One small error was found in STAN on one sample, in which a total of 7 trees had diameters that differed by a few centimeters from original datasheets; these were corrected to match original datasheets. Elevation data for samples in COLLINS, ELDO, KERN, and STAN were provided by collaborators and derived from LandFire DEM product (LANDFIRE 2022). Elevation was calculated for the centroid of each sample using bilinear interpolation of the four nearest pixel values.

3.4 Data Cleaning

To ensure consistency across datasets, several data cleaning steps were implemented. Records with incomplete data that prevented further analyses were removed, including: for STAN, 15 samples with no lot data; for KERN, 1 sample with no tree data; and for ELDO, 18 samples in which stocking or composition information was missing.

As the focus of this study is on forest structure, composition, and density, unforested areas with no live trees recorded were removed where they occurred, including: for STAN, 25 samples; and for ELDO, 3 samples. Other datasets did not contain records representing unforested areas, which were either not inventoried in original cruises or removed by collaborators in preparation of the datasets.

Samples with low stocking have been excluded from some analyses of these datasets in the past (Stephens et al. 2018, Collins et al. 2021), primarily to facilitate comparison to contemporary USFS Forest Inventory and Analysis data for which basal area <9m^2 ha^{-1} is considered nonforested and may indicate stands whose structure has been modified by timber harvest or other severe disturbances. However, in the context of these datasets, low stocking may be driven by a variety of factors, including edaphic conditions and frequent fire effects; as such, areas with BA <9m^2 ha^{-1} were retained to describe the full range of forested conditions across historical landscapes.

Samples showing evidence of disturbances that would modify forest structure from 'historical' conditions were removed, including 14 samples in KERN that showed evidence of logging. No other instances of logging or other disturbance were observed in the data.

Due to inconsistent recording of lodgepole in KIR, and the historically mixed-severity fire regime in lodgepole-dominated forests that differs from the frequent fire regimes otherwise common across datasets, 2,462 samples in which lodgepole pine made up >50% of the basal area were removed from KIR (Personal communication, R.K. Hagmann, Hagmann et al. 2019).

Finally, due to varying minimum diameter cutoffs across datasets, all tree data below the highest common diameter cutoff (30.5cm) was dropped. Datasets in which data were removed included ELDO, KIR, and STAN, as COLLINS and KERN already utilized a 30.5m lower diameter limit. While ELDO lacked tree-level data, the effects of removal of trees <30.5cm on sample-level variables were estimated from reported tree density in the 15.2-30.5cm range and an assumed mean diameter in the middle of that range (22.9cm). ELDO also contained species-specific BA estimates, which were decreased in proportion to the share of total BA made up by that species before the 30.5cm limit was applied. In KIR, 61 samples had all trees <30.5, and removal of those trees resulted in unforested conditions which required removal of the samples from the dataset.

These data cleaning steps resulted in retention of 269 samples for STAN, 379 samples for KERN, 631 samples for ELDO, 1,552 samples for COLLINS, and 18,018 samples for KIR.

3.5 Calculation of Structure and Composition Variables

Following data preparation and cleaning, forest structure variables were calculated at the sample scale, including quadratic mean diameter (QMD, cm), tree density (TPH, trees ha^{-1}), basal area (BA, m^2ha^{-1}), stand density index (SDI, metric units (see below), and relative stand density index (relative SDI, %). Species composition by BA was also calculated at the sample scale for all species recorded in each dataset, and pine and fir fraction were additionally calculated as the total of species composition values for *Pinus* and *Abies* species, respectively.

Variables were first calculated at the finest resolution present in the datasets: individual tree-level for STAN, KERN, and KIR; species-level averages within a sample for COLLINS; and averages for each sample for ELDO. Individual-tree data for STAN, KERN, and KIR were then averaged to the sample scale, converting to per hectare values using transect sizes. Species-level averages for COLLINS were averaged within each sample, and no aggregation was applied to ELDO as data were already at the sample scale.

SDI was calculated with two methods for later comparison: additive SDI (SDI_A), which utilizes information about individual trees (Equation 1), was calculated for STAN, KERN, and KIR, and traditional SDI (SDI_T), which utilizes stand-level average information (Equation 2), was calculated for all datasets (Reineke 1933, Long and Daniel 1990, Shaw 2000). All SDI values are in metric units, representing the density in trees ha^{-1} of an even-aged stand with QMD of 25.4cm experiencing the same level of density-dependent competitive pressure as the stand for which SDI is being calculated.

(Eq. 1) $SDI_A = \sum_i^N TPH_i * (\frac{DBH_i}{25.4})^{1.6}$, where N = number of trees sampled, TPH_i = trees ha^{-1} represented by the ith tree, and DBH_i = DBH of the ith tree (cm)

(Eq. 2) $SDI_T = TPH * (\frac{QMD}{25.4})^{1.6}$, where TPH = trees ha^{-1} and QMD = quadratic mean diameter (cm) of the sampled stand

The ratio of SDI_T to SDI_A (see Section 3.7 for details) was used to correct traditional SDI values for datasets without tree data in order to approximate additive SDI values. Specifically, the ratio SDI_A:SDI_T was calculated for all trees >12" in STAN, KERN, and KIR, and the mean of that ratio across samples (0.9805815) was multiplied by SDI_T to produce a corrected SDI (SDI_C) for ELDO and COLLINS. This represented a conservative correction, as the range of the ratio was 0.8789071-0.9999999 with median 0.9838545.

For each dataset, a 'best available' method of SDI calculation was chosen and used in analyses of SDI (referenced henceforth simply as SDI). For datasets with tree data (STAN, KERN, KIR) this was SDI_A, and for datasets with sample-average data (ELDO, COLLINS) this was SDI_C.

Maximum SDI values (SDI_{MAX}), representing the carrying capacity of a stand of a given species composition in units of SDI, were assigned to each sample. SDI_{MAX} values were assigned following North et al. (2022), by classifying each sample into one of three forest types and assigning a corresponding SDI_{MAX} value. Classification was based on the percent of stand total BA by genera: >50% *Pinus* was assigned "Pine Mixed Conifer" and SDI_{MAX} of 902; >50% *Abies* was assigned "Mesic Mixed Conifer" and SDI_{MAX} of 1359; and $\leq$50% *Pinus* and $\leq$50% *Abies* was assigned "Xeric Mixed Conifer" and SDI_{MAX} of 1112. Relative SDI was calculated for each sample from by dividing the 'best available' measure of SDI by the assigned SDI_{MAX} value. Relative SDI values were binned into four categories to descriptively characterize competitive dynamics, with 0%-25% representing no competition, 25%-35% representing little competition, with 'onset of competition' at 25%, 35%-60% representing moderate competition, with 'onset of full site occupancy' at 35%, and 60%-100% representing high competition, with 'imminent density-dependent mortality' beginning at 60% (Long 1985, Long and Shaw 2005).

While these methods of SDI calculation rely on numerous assumptions, including that SDI_{MAX} is invariant across site productivity and that the self-thinning line has a constant slope of approximately -1.6, SDI retains substantial utility as an integrated measure of growing space occupied by trees, and especially when combined with SDI_{MAX} values to estimate relative SDI, shows significant ability to accurately and simply characterize the competitive dynamics of forest stands (Chivhenge et al. 2024). As such, more detailed estimates of SDI_{MAX} are not used here, and the original linear slope of -1.6 is retained.

3.6 Structure and Composition Analysis

To characterize historical forests, structural and compositional variables were summarized within and across datasets to produce mean, standard deviation, and minimum-maximum ranges.

3.7 Assessing Effects of Analysis Choices

To assess the effect on estimates of forest structure and composition of inclusion of samples with low stocking (BA <$9m^2ha^{-1}$), mean values of structural variables were calculated within and datasets following removal of samples with BA <$9m^2ha^{-1}$ and were subtracted from mean values prior to removal to create a measure of change. Structural and compositional variables were additionally calculated for low BA samples to characterize the data that would be removed by a $9m^2ha^{-1}$ threshold. Furthermore, to assess the effect of raising the minimum diameter limit to 30.5cm (the highest limit in common across datasets), mean values of structural variables were calculated for datasets with data describing trees <30.5cm (ELDO, KIR, and STAN) utilizing both a 15.2cm and 30.5cm lower limit, with the difference taken as a measure of change.

To characterize differences produced by contrasting methods of SDI calculation in estimates of SDI in historical forests, the ratio of SDI_A:SDI_T was calculated at the sample level for datasets containing tree-level data (KERN, KIR, STAN), and summarized within and across datasets as mean, standard deviation, and minimum-maximum range (Ducey 2009). Additionally, as this ratio is known to be influenced primarily by the shape of the diameter distribution for which SDI values are being calculated, diameter distribution graphs were created for KERN, KIR, and STAN. Note that while ELDO contains tree density values by 15.2cm diameter classes, for which it is possible to calculate SDI_A by approximation of DBH as diameter class midpoint, use of wider diameter classes can introduce bias (Woodall et al. 2011), rendering inclusion of ELDO irrelevant here as SDI_A values were intended to represent a "most-accurate" assessment of stand density. Accordingly, a diameter distribution graph was not generated for ELDO.

3.8 Cluster Analysis

To explore broad-scale patterns of structure and composition across datasets, and to assess whether datasets reliably differed in key structure and composition variables, K-means clustering analysis was employed on all samples across datasets. For each sample, mean values of QMD, tree density, BA, relative SDI, and pine fraction were utilized. Since substantial correlation is known to exist between these variables, Principal Component Analysis (PCA) was used to simplify input variables into a smaller set of uncorrelated Principal Components (PC) that characterize the majority of variation in the input variables, with key PCs then used in K-means analysis. PCA was performed with the "rda" function from the "vegan" package on scaled and centered input variables, and key PCs were identified as the first *n* PCs that explained a majority of the variation in the data, and after which the proportion of variance explained by successive PCs sharply decreased (Gotelli and Ellison 2013, Oksanen et al. 2022). K-means analysis was performed with the "kmeans" function from the "stats" package, which generates a pre-determined number of clusters through an iterative process beginning with cluster centers at randomly chosen data points and re-assigning membership of data points to clusters until all data points in each cluster are closer to each other than to data points in other clusters (Gotelli and Ellison 2013, R Core Team 2022). To increase reliability of cluster identification, 25 sets of randomly chosen starting points were utilized with the most accurate classification chosen. Finally, the optimal number of clusters was identified as the cluster number for which increases in within-cluster similarity (reductions of sum of within-cluster sum of squares distances) produced by increasing cluster number sharply decreased from increases observed up to the optimal number of clusters.

Identified clusters were described by calculation of mean, standard deviation, and minimum-maximum ranges of structure and composition variables across samples within each cluster. Scatterplots of cluster membership in relation to key variables were generated to descriptively characterize clusters.

4. Results

4.1 Structure and Composition

4.1.1 Summary Across Datasets

Summaries across datasets showed historical forests with low density and BA, moderate to large trees, and low levels of competition (Table 2). Tree density averaged 31.8 trees ha^{-1} and BA averaged 11.5 m^2ha^{-1}, while QMD averaged 68.1 cm. SDI and relative SDI averaged 148.6 (metric units) and 15.8%, respectively. Relative SDI was consistently low across assigned SDI_{MAX} types, with 75^{th} percentiles near or below 35%, though with a smaller range in Pine Mixed-Conifer (Figure 1). Species composition was dominated by ponderosa pine, which was found on 97.1% of samples, followed by white fir (27.4%), sugar pine (21.0%), incense-cedar (17.0%), Douglas-fir (7.7%), red fir (4.5%), lodgepole pine (3.8%), and Jeffrey pine (2.2%). Other species were found on <1.5% of transects. While not all species were documented across datasets, the composition by BA of species that occurred in all datasets was also dominated by ponderosa pine (84.2%), followed by white fir (6.4%), sugar pine (4.1%), and incense-cedar (2.2%).

Minimum values reflected nearly-unforested areas, approaching zero for tree density, BA, SDI, and relative SDI (displayed as 0 in Table 2 due to rounding), while QMD had a minimum value of 21.9cm. The sample with the minimum QMD, which is below the lower diameter limit of 30.5cm, was in ELDO, for which removal of trees <30.5cm was approximated; this discrepancy is likely due to imprecision in the approximation procedure. Outside of this sample and one other in ELDO, minimum QMD across datasets was 30.5cm. Maximum values reflected samples with large trees (QMD 133.6cm), high stocking and density (BA 158.8 m^2ha^{-1}, tree density 354.6 trees ha^{-1}), and high levels of competition (SDI 1987.5 (metric), relative SDI 146.2%). The maximum value of QMD was observed on a sample in KIR with low tree density and BA, while maximum values of BA, tree density, SDI, and relative SDI were observed on a sample in ELDO. Across all datasets, relative SDI >100% was only seen on 18 samples in ELDO.

Substantial variability (standard deviation) relative to mean values was observed in BA, tree density, and SDI and relative SDI across datasets, with coefficients of variation (data not displayed) ranging from 68%-79%, while QMD showed much lower variability, with coefficient of variation averaging 15%.

4.1.2 Summary by Dataset

COLLINS included, on average, moderately sized trees (QMD 68.9cm), low BA (16.1 m^2ha^{-1}), low tree density (43.6 trees ha^{-1}), and low competition (SDI 209.2 (metric), relative SDI 22%) (Table 2). Species composition by BA was predominantly pine, with sugar pine and ponderosa pine making up 33.5% and 29.8% respectively, with a third major component of white fir (21.8%), with other species <8% each. Minimum values reflected areas with few trees, although QMD reached a minimum of only 52.2cm, while maximum values reflected samples with large trees (QMD 87.8cm), higher stocking (BA 42.9 m^2ha^{-1}), and higher tree density (123.3 trees ha^{-1}), while competition was maintained at comparatively moderate levels, with relative SDI reaching a maximum value of 62.6%. Compared to other datasets, COLLINS had relatively low QMD and intermediate BA, tree density, and competitive stress (relative SDI). Variability in structural variables was consistently lower in COLLINS than in other datasets.

ELDO included, on average, large trees (QMD 80.4cm), high stocking and tree density (BA 40.6 m^2ha^{-1}, 81.5 trees ha^{-1}), and moderate competition (SDI 494.4 (metric), relative SDI 42.5%). Species composition was relatively balanced between species and included more fir than pine, with white fir making up the largest component at 36.2%, followed by incense-cedar (17.2%), Jeffrey-pine (12.6%), sugar pine (10.5%), ponderosa pine (10.3%), and red fir (9.7%), with other species <2% each. Minimum values reflected areas with nearly no trees (BA 0.1 m^2ha^{-1}, tree density 1.2 trees ha^{-1}), and QMD (excluding samples for which QMD is <30.5cm and may be inaccurate) reached a minimum of 37.7cm, approaching the limit of 30.5cm imposed by this analysis. Maximum values represented dense stands experiencing very high competition (relative SDI 146.2%). Compared to other datasets, ELDO has higher stocking, density, and competition, with maximum values across datasets of BA, tree density, SDI and relative SDI occurring in ELDO. Variability in structural variables was also consistently higher in ELDO than in other datasets.

KERN included, on average, moderately sized trees (QMD 73.7cm), intermediate stocking and tree density (BA 22 m^2ha^{-1}, 54.5 trees ha^{-1}), and low competition (SDI 273.7 (metric), relative SDI 25.4%). Species composition was relatively balanced between species and included 37.4% ponderosa pine, 29.3% incense-cedar, 24.7% white fir, and 8.5% sugar pine. Similar to other datasets, minimum values reflected areas with few trees (2.5 trees ha^{-1}) while QMD reached a minimum of only 42.3cm. Maximum values reflected only moderately dense stands (170.6 trees ha^{-1}) experiencing high but not extreme competition (relative SDI 78.8%). Compared to other datasets, KERN shows higher stocking and density than COLLINS, KIR, and STAN, but lower values than ELDO; however, KERN is closer to COLLINS, KIR, and STAN than ELDO in these metrics. Variability of structural variables follows the same pattern, with KERN consistently showing second-highest variability behind ELDO.

KIR included, on average, moderately sized trees (QMD 67.6cm), and very low BA (9.8 m^2ha^{-1}), tree density (28.6 trees ha^{-1}), and competition (SDI 128.5 (metric), relative SDI 14.1%). Species composition was nearly exclusively pine, with ponderosa pine making up 93.2% by BA, although white fir was the next-most abundant species at 3.7%, followed by sugar pine (1.2%), with other species making up <1% each. Minimum values approached zero, representing areas with extremely few trees, while minimum QMD reached 30.5cm, the value imposed by this analysis; maximum values represented moderately dense conditions (294.1 trees ha^{-1}) experiencing high but not extreme levels of competition (relative SDI 83.3%). Compared to other datasets, KIR has smaller trees, lower stocking and density, and lower levels of competition. Variability in structural metrics was consistently lower than in ELDO and KERN and higher than COLLINS, showing similarity to variability observed in STAN.

STAN included, on average, moderately sized trees (QMD 74.4cm), low BA (15.6 m^{2}ha^{-1}), low tree density (35.4 trees ha^{-1}), and low competition (SDI 192 (metric), relative SDI 20%). Species composition was majority pine, with ponderosa pine and sugar pine making up 49.8% and 8.8% respectively, followed by incense-cedar (30.8%), Douglas-fir (6.5%), and white fir (3.5%), with other species <1% each. Minimum values represented areas of low stocking (BA 0.5 m^{2}ha^{-1}) and tree density (1.8 trees ha^{-1}), but minimum QMD only reached 47.9cm. Maximum values represented areas with only moderate levels of stocking (BA 38.5 m^{2}ha^{-1}), tree density (91.4 trees ha^{-1}), and competition (relative SDI 49.2%). Compared to other datasets, STAN shows intermediate tree sizes, moderately low stocking, low tree densities, and low competition, with only KIR having lower densities and competition. Maximum values in STAN are consistently much lower than in other datasets, in some cases falling near (tree density, relative SDI) or below (BA, SDI) mean values observed in ELDO. Variability in STAN is similar to KIR, being lower than ELDO and KERN but higher than COLLINS.

Across datasets, relative SDI was consistently low, with the vast majority of data showing little to no competitions (<35%) (Figure 2). ELDO showed a wider range of stocking, but the 75th percentile was still below the initiation of high competition at 60% relative SDI.

4.2 Analysis Choices

4.2.1 Low Stocking Samples

Samples with low BA (<9m^{2}ha^{-1}) made up significant components of KIR (47.5% of samples), STAN (20.4%), and KERN (15%), and small components of COLLINS (5.3%) and ELDO (2.7%) (Table 3). Across datasets, samples with low BA made up 42.1% of total samples. As expected, mean values of structural variables were generally low, with BA of 4.5 m^{2}ha^{-1}, tree density of 14 trees ha^{-1}, SDI (metric) of 59.5, and relative SDI of 6.5%. However, mean QMD of 66.6cm was relatively high, showing that low BA samples generally consisted of few large trees. Interestingly, species composition showed that low BA samples contained a greater fraction of pines, within each dataset and averaged across datasets, than for samples of all BA (Table 2).

Removal of low BA samples primarily resulted in increases of varying size in structural variables averaged across samples, both within and across datasets (Table 4). QMD increased by 1.2cm, BA by 5.1 m^{2}ha^{-1}, tree density by 13 trees ha^{-1}, SDI by 65.1 (metric), and relative SDI by 6.8%. Within datasets, magnitude of changes generally increased with the percent of low BA samples within that dataset, with the largest changes observed in KIR, followed by STAN and KERN; however, despite COLLINS having a higher percentage of low BA samples than ELDO, changes were less pronounced in COLLINS than ELDO. The only decrease in mean value of a structural variable produced by removing low BA samples was a decrease in QMD of 0.2cm in KERN, corresponding to a higher QMD in low BA samples (Table 3) than for samples of all BA (Table 2).

4.2.2 Minimum Diameter Limit

Trees with diameters 15.2-30.5cm made up an average of 42.0% percent of tree density in the 3 datasets with data in that range (STAN, KIR, ELDO) (Table 5, percent calculated as *Change in Tree Density / Tree Density at 15.2cm * 100*). However, the percent varied by dataset, with KIR having the highest proportion (42.9%), followed by ELDO (35.1%) and STAN (25.9%).

Raising the minimum diameter limit from 15.2cm to 30.5cm produced differences in forest structure variables averaged across samples. QMD increased by 13.9cm, while BA decreased by 0.9 m^2ha^{-1}, and tree density decreased by 22.1 trees ha^{-1}. These changes resulted in decreases in SDI and relative SDI by 19.4 (metric units) and 2.1%, respectively (Table 5). Changes within each dataset generally followed the pattern of mean changes across datasets; however, mean changes across datasets are highly similar to changes observed in KIR due to large sample size in KIR, obscuring notable differences such as the relatively larger decrease in tree density in ELDO (44.1 trees ha^{-1}). Importantly, changing the minimum diameter limit resulted in the exclusion of 61 samples in KIR that became unforested when considering only trees $\geq$ 30.5cm; these samples were not included in comparisons of structure and composition metrics to avoid inclusion of unforested areas.

Few substantial changes ($\geq$ 1% change) to species composition, as measured by BA, were observed. In KIR, ponderosa pine increased in mean abundance from 91.1% to 93.2%, a change of +2.1%, while lodgepole pine decreased in mean abundance from 2.4% to 0.3%, a change of -2.2% (values rounded).

4.2.3 SDI Calculation Method

For datasets with tree-level data (STAN, KERN, KIR), comparison of additive and traditional SDI values calculated at the sample scale (Equations 1 and 2, respectively) revealed little difference, with a mean ratio of SDI_A:SDI_T of 0.981 (Table 6). Variation from this mean did occur by dataset; the large sample size in KIR (18,018 samples) meant that the mean in KIR (0.981) obscured the mean in STAN (0.966) and KERN (0.962) in calculation of overall mean ratio. Variation in the ratio was small, with standard deviation averaging 0.013; however, the range of the ratio was relatively wide, with minimum values between 0.879-0.910, and maximum values between 0.996-1. Diameter distributions for these datasets (Figure 3) show shapes reasonably similar to a normal distribution for trees >30.5cm; however, as this ratio is known to be influenced by the minimum diameter cutoff chosen, the ratio was additionally calculated for datasets with tree-level data in the 15.2-30.5cm range (STAN, KIR), producing a mean ratio of 0.94 (standard deviation 0.028, range 0.792-1.00). Diameter distributions with a 15.2cm lower diameter limit show strongly right-skewed distributions (Figure 3).

4.3 Clustering Analysis

Principal component analysis identified 3 key principal components explaining 98.2% of the variation of the input data. PC1 explained 62.4% of the variance was strongly positively related to tree density, BA, and relative SDI, and moderately negatively related to pine fraction. PC2 explained 21.6% of the variance and was strongly negatively related to QMD and weakly positively related to tree density. PC3 explained 14.2% of the variance, and was strongly negatively related to pine fraction, and weakly negatively related to relative SDI and tree density.

K-means clustering on the 3 key PCs identified 4 distinct clusters with unique structural and compositional characteristics (Table 7). Clusters contained relatively comparable numbers of samples (6645, 6044, 2677, and 5843 samples in Clusters 1-4, respectively), and were differentiated first by pine and fir fraction, with Cluster 3 showing lower mean pine fraction (41.9%) and higher mean fir fraction (37.3%) than Clusters 1, 2, and 4 (pine fraction 96.6%, 95%, 97.6%; fir fraction 2%, 3.6%, 1.6%) (Figure 4, A). Within the high-pine group, clusters were differentiated primarily by level of stocking and competition (BA, relative SDI), with lower stocking and competition in Cluster 1 (5.7 m^2ha^{-1}, 8.7%) and Cluster 4 (7.9 m^2ha^{-1}, 10.9%) and higher stocking and competition in Cluster 2 (17.2 m^2ha^{-1}, 24.7%). Additionally, in high-pine, low-stocking Clusters 1 and 4, minimum pine fraction and maximum fir fraction remained >50% and <50%, respectively, while in high-pine, high-stocking Cluster 2, minimum pine fraction and maximum fir fraction extended to 2.5% and 97.5%, respectively. High-pine, low-stocking Clusters 1 and 4 were differentiated by lower QMD in Cluster 1 (60.1cm) than Cluster 4 (78.2cm) (Figure 4, B). Comparatively, low-pine Cluster 3 showed high levels of stocking and competition and moderately high QMD (20.5 m^2ha^{-1}, 23.7%, 69.8cm) comparable to high-pine Cluster 2 (17.2 m^2ha^{-1}, 24.7%, 67.5cm). Cumulatively, Cluster 1 represents samples with moderate QMD, low stocking and competition, and high pine and low fir fractions; Cluster 2 represents samples with moderately high QMD, higher stocking and competition, and a wide range of pine and fir fraction values that nevertheless have high and low means, respectively; Cluster 3 represents samples with moderately high QMD, higher stocking and competition, and lower pine and higher fir fractions; and Cluster 4 represents samples with high QMD, low stocking and competition, and high pine and low fir fraction. Variability (standard deviation) of structure and composition variables was consistently highest in Cluster 3, lowest in Cluster 1 and 4, and intermediate in Cluster 2.

Datasets from each site were distributed across clusters, although the share of a dataset contained in each cluster varied (Table 8). However, membership in Cluster 3, a high-stocking cluster that contained ~2x lower pine fraction and ~10-20x higher fir fraction than other clusters, differentiated datasets in California (COLLINS, ELDO, KERN, and STAN), which showed high membership in Cluster 3 (>50% of samples), from the southern Oregon dataset (KIR), which showed low membership in Cluster 3 (<5% of samples). Nevertheless, the California datasets all showed substantial membership in Clusters 1, 2, and 4 (between 22-42% of samples across Clusters 1, 2, and 4) alongside KIR, which had ~33% of samples in each of Clusters 1, 2, and 4.

5. Discussion

Synthesis of this set of robust historical forest inventories generated similar estimates of historical forest structure and composition in ponderosa pine and mixed-conifer forests in California and Southern Oregon as reported from a variety of historical data sources (Safford and Stevens 2017, Bohlman et al. 2021). Historical forests were characterized by relatively low densities of large trees, contained high amounts of pine relative to other species, and experienced little to no competition. While this study did not present new data, the synthesis of previously uncombined datasets helps characterize regional-scale patterns, documenting substantial variability in forest structure and composition likely produced by the interaction of site-specific environmental and topographical factors with fire regimes produced by Indigenous cultural fire use and lightning ignitions (SUGIHARA et al. 2018, Long et al. 2021, Greenler et al. 2024).

These results build on evidence documenting low levels of stand-average competition in these historical forests, suggesting that the low densities produced by frequent fire facilitated relatively high levels of individual tree vigor (North et al. 2022, Nepal et al. 2023). Sustained high individual tree vigor, in turn, is correlated with resilience of forest stands, or retention of characteristic ecosystem composition, structure, and function through disturbance events common to that ecosystem type at long (300-400 year) time scales (Franklin et al. 1987, Das et al. 2016, Cailleret et al. 2019). Comparison of resilience across forest types through calculation of relative SDI relies on accurate assignment of an SDI_{MAX} value, which was originally assumed to be constant across site factors for a given forest type, but has been shown to vary with site productivity (including climate, topography, and soils) and species composition within the forest type (Chivhenge et al. 2024). Accordingly, assignment of one of three SDI_{MAX} values to each sample in these datasets is an oversimplification of historical competitive dynamics. However, absolute values of SDI for samples are exceedingly low, such that use of site-specific SDI_{MAX} values for ponderosa pine and mixed-conifer forests across the study sites, generally between 800-2000 (metric), would produce relative SDI values with the same general trend of little to no competition across samples (Kimsey et al. 2019, Woodall and Weiskittel 2021 Supplementary Figure 1).

While datasets generally all describe low-density forests with large trees, differences between datasets exist that may be partially accounted for by site factors such as elevation and soils. For example, ELDO consistently showed the highest levels of density, stocking, tree size, and competition, likely attributable to the greater water availability and lower evaporative demand associated with the higher elevation site and higher precipitation levels represented in ELDO (Stephenson 1998). However, mean elevation calculated from LandFire elevation data for ELDO was only slightly higher than for KERN, suggesting that elevation may be an incomplete explanation and other factors, such as soil depth and texture, may support higher water availability and therefore productivity in ELDO (O'Geen et al. 2007). Similarly, KIR consistently showed the lowest levels of density, stocking, tree size, and competition, and shows both lower elevation than other datasets, and distinct soil types. Lower elevation in KIR is accompanied by relatively low precipitation (60-70cm yr^{-1}); while reported precipitation is much lower for KERN (32cm yr^{-1}), that value may not accurately characterize precipitation in KERN, as North et al. (2022) described precipitation across KERN and STAN as having a minimum of 83cm. As such, KIR likely contains the lowest levels of precipitation of all datasets. Additional reductions in tree density and size in KIR may derive from the influence of soils, the majority of which are pumice soils derived from the eruption of Mount Mazama approximately 7,500 years before present, which differ significantly from the mineral soils underlying the other datasets (Carlson 1979, Egan et al. 2015). Pumice soils have poor heat storage and transfer, resulting in more frequent frosts and frost-heaving throughout the year, and can show rapid increases in shallow soil temperatures in extremes of air temperature, both of which increase seedling mortality (Carlson 1979). Finally, the presence of 'outlier' values in ELDO and KIR may be attributable to characteristics of the historical fire regime in those areas. While not directly documented, the fir-dominated mixed-conifer forest represented in ELDO likely experienced less frequent fire than other datasets, as the mid- to low-elevation red fir forest into which ELDO transitions experienced fire return intervals from 27-51 years (Meyer and North 2019). Less frequent fire could have allowed establishment of dense regeneration clumps which subsequently reduce mortality from fire by modifying microclimate conditions such that fire intensity is reduced, as well as by outer trees providing direct fire protection for inner trees (Ma et al. 2010, Blomdahl et al. 2019). Similarly, while estimated fire return intervals are similar across datasets, KIR may have experienced more frequent very low-severity fires that were not recorded in FRI estimates but that may still reduce density, as there are few topographic barriers to fire spread in the Cascade Range in southern Oregon, potentially allowing widespread low-severity fire to travel across the landscape (Merschel et al. 2018).

Despite these distinctions between datasets, shared structural and compositional types were identified by PCA across datasets, suggesting a top-down influence of frequent fire in homogenizing forest conditions at the landscape scale (Chamberlain et al. 2023). Furthermore, variability of structural metrics was generally high except for QMD, supporting the notion that frequent fire interacted with site-specific factors to vary tree densities, but that average tree size was consistently large across the historical landscape (Safford and Stevens 2017).

Analyses of common data preparation and cleaning choices showed generally expected effects on stand structural and compositional metrics. Historical datasets excluding trees <30.5cm DBH likely underestimate tree density, BA, and relative SDI by approximately 26-43%, 3-8%, and 5-13% respectively, and overestimate QMD by 14-26%, relative to data including trees 15.2cm-30.5cm (values calculated from Table 5). Little influence on estimates of species composition was found, except in KIR, where increase and decrease in ponderosa and lodgepole pine, respectively, suggest larger proportions of those species occurring above and below the 30.5cm cutoff, respectively. Effects on these metrics of inclusion of trees <15.2cm is not represented here. Exclusion of low-stocking samples likely overestimates tree density, BA, relative SDI, and QMD by approximately 2-47%, 2-50%, 2-49%, and 0-2%, respectively (values calculated from Table 4). The relatively minor changes in QMD produced by removal of low-stocking samples may suggest that areas of low stocking are characterized more by low tree densities than by smaller trees. Characterization of the ratio of additive to traditional SDI was influenced, as expected, by the shape of the diameter distribution, which is itself strongly determined by the chosen minimum diameter limit, and showed that use of traditional SDI likely overestimates SDI by 2-4% (Table 6). Future investigations of the influence of analysis method choices on historical SDI could characterize the effect of increasing spatial scales on SDI estimates, as evidence suggests that calculation of structural metrics at larger scales in similar forest types may reduce estimates of variability across the landscape compared to calculation at smaller scales (Ritchie 2016). If scale also affects estimates of mean values, this could mean that estimates of SDI calculated from single metrics reported for individual studies or forest types within studies may less accurately characterize historical forest dynamics.

Finally, differences between values described here and elsewhere for each dataset likely originate from differing choices of which data to include; for example, low BA samples were removed before analysis ELDO and COLLINS in their respective papers but included here, while minimum diameters were set at 15.2cm for ELDO, KIR, and STAN in their respective papers but set at 30.5cm here (Hagmann et al. 2013, Collins et al. 2015, 2021, Stephens et al. 2018). Additionally, a much larger number of samples is described here for KIR than in other publications drawing on the dataset, with 18,018 samples (each containing 1-4 transects) described here, and other publications reporting 3068 transects (Hagmann et al. 2013), 1036 transects (Hagmann et al. 2017), and 9,374 transects (Hagmann et al. 2019). The inclusion of such a large amount of previously undescribed data likely accounts for differences between values reported here and in other publications drawing on this dataset.

5.1 Management Implications

The data assembled here support the contention of North et al. (2022) that if managers seek to create and maintain the high tree vigor associated with open-grown conditions seen in the historical period, the density management zone would need to be shifted downward such that 35% relative SDI represents the maximum management threshold, rather than the minimum, as is common in traditional density management systems (Drew and Flewelling 1979). Approximately 75-95% of the samples analyzed here fell below the 35% relative SDI threshold (Figure 1). However, the inclusion of samples with low stocking likely means that these estimates include areas where low density is due to edaphic conditions, such as rock outcroppings, very poor soils, or other conditions which contemporary forest managers would exclude from the forested areas under active management. Almost certainly, however, these low stocking samples also included areas of otherwise productive forestland that supported low densities due to the effects of frequent fire. As these low stocking areas made up variable proportions of total forested area across the forest types represented in these datasets (2.7%-42.1%), managers could approach creation of these very low stocking levels as part of an approach that tailors prescriptions to local site conditions, including elevation, topographic position, and forest type (North et al. 2009). The wide range of conditions documented across datasets, including a wide range of species compositions with variable dominance of pine, further supports the use of variability and site-specific management goals across landscapes, with heterogeneity across scales supporting a variety of ecosystem processes and bolstering ecosystem resilience (North et al. 2009, Palik et al. 2020, Ziegler et al. 2021).

6. Author Statements:

Acknowledgements: We appreciate the assistance of Alexis Bernal in generating and sharing elevational data for datasets in California.

Competing interests: The authors declare there are no competing interests.

Author contributions:

Funding statement:

Data availability:

7. Tables

Table 1. Sampling methods used across datasets. Quarter-Quarter Section is abbreviated to QQ.

Dataset	COLLINS	ELDO	KERN	KIR	STAN
Year	1924	1923-36*	1911	1914-1922	1911

Data Recorded	For each QQ, for each species: tree counts and mean diameter	For each QQ: tallies by diameter class (15.2cm classes); BA by species	For each QQ, for each species: diameter class tallies (5.1cm classes)	For each Sample, for each species: diameter class tallies	For each QQ by species: diameter class tallies (15.2-30.5cm then 5.1cm classes)
Spatial Scale of Data	QQ	QQ	QQ	Quarter; Half-QQ	QQ
Resolution of Data	Averages	Averages	Individual Tree	Individual Tree	Individual Tree
Sampling Layout	4 transects per QQ	Majority 2 transects per QQ	1 transect per QQ	1 transect per QQ combined to 4 per Quarter; 1 transect per half-QQ	1 transect per QQ
Sampling Intensity	40% per QQ	Majority 10% per QQ	5% per QQ	10% per Q; 20% per half-QQ	10% per QQ
Minimum Diameter (cm)	30.5	15.2	30.5	15.2	15.2
Species Included	Conifers	All Species	Conifers	Conifers	Conifers

**Data in ELDO adjusted to represent conditions in 1923 (Stephens et al. 2018)*

Table 2. Summary of structural and composition variables within and across datasets. Values are reported as mean ± standard deviation (minimum – maximum).

	COLLINS	**ELDO**	**KERN**	**KIR**	**STAN**	**Total**
QMD (cm)	68.9 ± 5.6 (52.2 - 87.8)	80.4 ± 11.4 (21.9 - 129.6)	73.7 ± 10.9 (42.3 - 109.9)	67.6 ± 10.4 (30.5 - 133.6)	74.4 ± 9.6 (47.9 - 106.6)	68.3 ± 10.5 (21.9 - 133.6)
BA (m^2ha^{-1})	16.1 ± 4.6 (0.5 - 42.9)	40.6 ± 18.4 (0.1 - 158.8)	22 ± 12.2 (1.1 - 59.7)	9.8 ± 6.4 (0 - 59.2)	15.6 ± 7.6 (0.5 - 38.5)	11.5 ± 9.1 (0 - 158.8)
Tree density (trees ha^{-1})	43.6 ± 12.4 (2 - 123.3)	81.5 ± 38.9 (1.2 - 354.6)	54.5 ± 33 (2.5 - 170.6)	28.6 ± 20.3 (0 - 294.1)	35.4 ± 15.6 (1.8 - 91.4)	31.9 ± 23.2 (0 - 354.6)
SDI (metric)	209.2 ± 58.2 (7.5 - 564.6)	494.4 ± 222.3 (2.2 - 1987.5)	273.7 ± 151.9 (15.4 - 710.4)	128.5 ± 82.5 (0.1 - 849.1)	192 ± 90 (6.9 - 449.5)	149 ± 112.7 (0.1 - 1987.5)
Relative SDI (%)	22 ± 6.1 (0.7 - 62.6)	42.5 ± 18.5 (0.2 - 146.2)	25.4 ± 13.3 (1.7 - 78.8)	14.1 ± 9 (0 - 83.3)	20 ± 9.5 (0.6 - 49.2)	15.8 ± 10.8 (0 - 146.2)
% Pine (*Pinus*)	63.3 ± 16.4 (2.7 - 100)	34.5 ± 20.8 (0 - 100)	45.9 ± 31.1 (0 - 100)	94.9 ± 14.2 (0 - 100)	58.7 ± 19.8 (0 - 100)	89.4 ± 21.1 (0 - 100)
% Fir (*Abies*)	24.1 ± 12.7 (0 - 97.3)	46 ± 26.8 (0 - 100)	24.7 ± 22.9 (0 - 100)	3.7 ± 12.2 (0 - 100)	4 ± 14.1 (0 - 100)	6.9 ± 16.1 (0 - 100)
% *Pinus ponderosa*	29.8 ± 23 (0 - 100)	10.3 ± 15.5 (0 - 82)	37.4 ± 32.5 (0 - 100)	93.2 ± 16.5 (0 - 100)	49.8 ± 22.1 (0 - 100)	84.4 ± 28.6 (0 - 100)
% *Abies concolor*	21.8 ± 12.3 (0 - 77.8)	36.2 ± 20.4 (0 - 96.9)	24.7 ± 22.9 (0 - 100)	3.7 ± 12.1 (0 - 100)	3.5 ± 11.3 (0 - 82.4)	6.4 ± 14.8 (0 - 100)
% *Pinus lambertiana*	33.5 ± 19.1 (0 - 89.4)	10.5 ± 11.4 (0 - 100)	8.5 ± 10.3 (0 - 48.4)	1.2 ± 4.9 (0 - 94.8)	8.8 ± 13.3 (0 - 100)	4.1 ± 11.4 (0 - 100)
% *Calocedrus decurrens*	5.5 ± 6.1 (0 - 37)	17.2 ± 14 (0 - 65.3)	29.3 ± 22.7 (0 - 90.9)	0.5 ± 2.4 (0 - 100)	30.8 ± 16.6 (0 - 75.9)	2.3 ± 7.8 (0 - 100)
% *Pseudotsuga menziesii*	7.1 ± 10.1 (0 - 45.1)	0.6 ± 3.3 (0 - 39)		0.5 ± 3.3 (0 - 100)	6.5 ± 15.1 (0 - 94.5)	1.1 ± 4.9 (0 - 100)

% *Pinus sabiniana*					0.1 ± 1 (0 - 12.8)	0.1 ± 1 (0 - 12.8)
% *Pinus monticola*		0 ± 0.5 (0 - 9.3)		0.2 ± 2.5 (0 - 100)	0 ± 0.2 (0 - 2.8)	0.2 ± 2.4 (0 - 100)
% *Abies magnifica*	2.4 ± 4.6 (0 - 37.4)	9.7 ± 19.1 (0 - 100)			0.5 ± 5.1 (0 - 59)	4.1 ± 11 (0 - 100)
% *Juniperus occidentalis*		0 ± 0.5 (0 - 10.5)		0.1 ± 1.7 (0 - 79.3)		0.1 ± 1.7 (0 - 79.3)
% *Pinus contorta* var. *murrayana*		1.1 ± 5 (0 - 64.2)		0.3 ± 2.2 (0 - 50.1)		0.3 ± 2.3 (0 - 64.2)
% *Pinus jeffreyi*		12.6 ± 14.9 (0 - 100)				12.6 ± 14.9 (0 - 100)
% *Quercus chrysolepis*		0 ± 0.7 (0 - 15.4)				0 ± 0.7 (0 - 15.4)
% *Quercus kelloggii*		1.7 ± 7.4 (0 - 79.3)				1.7 ± 7.4 (0 - 79.3)
% *Abies procera*				0 ± 1.4 (0 - 95.1)		0 ± 1.4 (0 - 95.1)
% *Tsuga mertensiana*				0.3 ± 4.7 (0 - 100)		0.3 ± 4.7 (0 - 100)
% *Pinus albicaulis*				0 ± 1 (0 - 100)		0 ± 1 (0 - 100)

Table 3. Characteristics of low BA samples (<9m^{2}ha^{-1}) within and across datasets. Structural and composition variables reported are mean values.

	Count of Samples	Percent of Dataset	QMD (cm)	BA (m^{2}ha^{-1})	Tree Density (trees ha^{-1})	SDI (metric)	Relative SDI (%)	% *Pinus*
COLLINS	83	5.3	68.3	7	19.5	91	9.6	66.6
ELDO	17	2.7	70.2	5.4	18.7	70.6	6.7	47.5
KERN	57	15	74.8	6	14.6	73.6	7.8	77.3
KIR	8566	47.5	66.5	4.4	13.9	58.9	6.5	96.8
STAN	55	20.4	68.1	6	16.9	77	7.9	59.9
Total	8778	42.1	66.6	4.5	14	59.5	6.5	96

Table 4. Effects of removing low BA samples (<9m^{2}ha^{-1}) on mean structural variables within and across datasets.

Low BA Exclusion	**QMD (cm)**	**BA (m^{2}ha^{-1})**	**Tree Density (trees ha^{-1})**	**SDI (metric)**	**Relative SDI (%)**
			COLLINS		
Included	68.9	16.1	43.6	209.2	22
Excluded	68.9	16.6	45	215.9	22.7
Change	0	0.5	1.4	6.7	0.7
			ELDO		
Included	80.4	40.6	81.5	494.4	42.5
Excluded	80.7	41.5	83.3	506.2	43.5
Change	0.3	1	1.7	11.7	1

KERN					
Included	73.7	22	54.5	273.7	25.4
Excluded	73.5	24.8	61.5	309.1	28.5
Change	-0.2	2.8	7.1	35.4	3.1
KIR					
Included	67.6	9.8	28.6	128.5	14.1
Excluded	68.5	14.7	41.9	191.5	21
Change	0.9	4.9	13.3	63	6.9
STAN					
Included	74.4	15.6	35.4	192	20
Excluded	76	18	40.2	221.5	23.1
Change	1.6	2.5	4.8	29.6	3.1
Total					
Included	68.3	11.5	31.9	149	15.8
Excluded	69.4	16.6	44.9	214.1	22.6
Change	1.2	5.1	13	65.1	6.8

Table 5. Effects on mean structural variables of raising minimum diameter limit from 15.2cm to 30.5cm within and across datasets (ELDO, KIR, STAN) containing data in the 15.2-30.5cm range.

Minimum Diameter	**QMD (cm)**	**BA (m^2ha^{-1})**	**Tree Density (trees ha^{-1})**	**SDI (metric)**	**Relative SDI (%)**
ELDO					
15.2cm	68.4	42.4	125.6	531.5	45.7
30.5cm	80.4	40.6	81.5	494.4	42.5
Change	12	-1.8	-44.1	-37.1	-3.2
KIR					
15.2cm	53.5	10.7	50.1	147.4	16.2
30.5cm	67.6	9.8	28.6	128.5	14.1
Change	14	-0.9	-21.5	-18.9	-2.1
STAN					
15.2cm	65.2	16	47.8	201.2	20.9
30.5cm	74.4	15.6	35.4	192	20
Change	9.1	-0.4	-12.4	-9.2	-1
Total					
15.2cm	54.2	11.9	52.6	161	17.2
30.5cm	68.1	10.9	30.4	141.6	15.1
Change	13.9	-0.9	-22.1	-19.4	-2.1

Table 6. Summary metrics for ratio of additive SDI to traditional SDI at the sample level within and across datasets containing tree-level data (KERN, KIR, STAN). Following other analyses, only trees ≥30.5cm were considered.

Dataset	Mean	St. Dev.	Min	Max
KERN	0.962	0.014	0.879	1
KIR	0.981	0.012	0.881	1
STAN	0.966	0.012	0.91	0.996
Total	0.981	0.013	0.879	1

Table 7. Summary of structural and composition variables within 4 clusters identified by K-means analysis. Values are reported as mean ± standard deviation (minimum – maximum).

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
QMD (cm)	60.1 ± 7.3 (30.5 - 70.5)	67.5 ± 7.8 (34.5 - 111.7)	69.8 ± 10 (21.9 - 107.5)	78.2 ± 7.5 (68.8 - 133.6)
BA ($m^2 ha^{-1}$)	5.7 ± 3.5 (0 - 16.4)	17.2 ± 7.3 (10 - 158.8)	20.5 ± 13.2 (0.1 - 87)	7.9 ± 5.1 (0 - 48.4)
Tree density (trees ha^{-1})	20.6 ± 13.7 (0 - 100.7)	48.9 ± 19.3 (26.9 - 354.6)	52.1 ± 28.7 (0.4 - 294.1)	16.8 ± 10.3 (0 - 51.9)
SDI (metric)	78.2 ± 47.2 (0.1 - 230.7)	224 ± 88.6 (130.6 - 1987.5)	260 ± 158.3 (1.3 - 1017)	98 ± 61.9 (0.2 - 523)
Relative SDI (%)	8.7 ± 5.2 (0 - 25.6)	24.7 ± 8.7 (14.5 - 146.2)	23.7 ± 12.4 (0.1 - 78.9)	10.9 ± 6.9 (0 - 58)
% Pine (*Pinus*)	96.6 ± 7.7 (59.5 - 100)	95 ± 10.5 (2.5 - 100)	41.9 ± 19.3 (0 - 71.8)	97.6 ± 7 (53.5 - 100)
% Fir (*Abies*)	2 ± 5.8 (0 - 40.5)	3.6 ± 7.9 (0 - 97.5)	37.3 ± 25.7 (0 - 100)	1.6 ± 5.5 (0 - 44.6)
% *Pinus ponderosa*	94.5 ± 13.5 (0 - 100)	90.9 ± 18.4 (0 - 100)	24.2 ± 21 (0 - 70.7)	94.5 ± 16.4 (0 - 100)
% *Abies concolor*	1.9 ± 5.7 (0 - 40.5)	3.5 ± 7.5 (0 - 54.8)	34.2 ± 23.6 (0 - 100)	1.5 ± 5.1 (0 - 39)
% *Pinus lambertiana*	1.8 ± 7.6 (0 - 81.1)	3.6 ± 10.1 (0 - 94.8)	14.1 ± 16.4 (0 - 66.7)	2.7 ± 11 (0 - 100)
% *Calocedrus decurrens*	0.9 ± 3.2 (0 - 36.8)	1 ± 4.1 (0 - 43.8)	12.1 ± 16.7 (0 - 100)	0.5 ± 2.9 (0 - 39.3)
% *Pseudotsuga menziesii*	0.3 ± 2.1 (0 - 31.1)	0.4 ± 2.3 (0 - 36.9)	6.8 ± 11.6 (0 - 100)	0.3 ± 2 (0 - 30.8)
% *Pinus sabiniana*	0.6 ± 2 (0 - 9.4)	0 ± 0 (0 - 0)	0.1 ± 1 (0 - 12.8)	0 ± 0 (0 - 0)
% *Pinus monticola*	0.1 ± 2.3 (0 - 100)	0.1 ± 0.7 (0 - 22)	1.6 ± 6.7 (0 - 57.7)	0 ± 0.8 (0 - 42.5)
% *Abies magnifica*	2.5 ± 4.1 (0 - 22.2)	1.8 ± 4.4 (0 - 67.4)	5 ± 13.2 (0 - 100)	3.5 ± 5.3 (0 - 28.7)
% *Juniperus occidentalis*	0.1 ± 1.5 (0 - 40.4)	0 ± 0 (0 - 1.4)	0.5 ± 5.3 (0 - 79.3)	0 ± 0.2 (0 - 11)
% *Pinus contorta* var. *murrayana*	0.2 ± 2.6 (0 - 50.1)	0.2 ± 2.1 (0 - 64.2)	0.9 ± 3.9 (0 - 44.6)	0.3 ± 1.6 (0 - 30.4)
% *Pinus jeffreyi*	79 ± 29.7 (58 - 100)	12.6 ± 14 (0 - 58.2)	11.4 ± 12.8 (0 - 68.5)	26.6 ± 26.2 (0 - 84.2)
% *Quercus chrysolepis*	0 ± 0 (0 - 0)	0 ± 0 (0 - 0)	0.1 ± 0.8 (0 - 15.4)	0 ± 0 (0 - 0)
% *Quercus kelloggii*	0 ± 0 (0 - 0)	0.6 ± 2.4 (0 - 19.5)	2.1 ± 8.3 (0 - 79.3)	0.2 ± 0.9 (0 - 5.4)
% *Abies procera*	0 ± 0 (0 - 0)	0 ± 0 (0 - 2.9)	0.6 ± 6.6 (0 - 95.1)	0 ± 0 (0 - 0)
% *Tsuga mertensiana*	0 ± 0.3 (0 - 21.5)	0 ± 0.1 (0 - 9.9)	6 ± 20.7 (0 - 100)	0 ± 0 (0 - 0)
% *Pinus albicaulis*	0 ± 1.4 (0 - 100)	0 ± 0 (0 - 0)	0.3 ± 2.4 (0 - 32.6)	0 ± 0 (0 - 0)

Table 8. Membership of datasets in 4 clusters identified by K-means analysis. Due to large number of samples, KIR composes >90% of Clusters 1, 2, and 4, and ~33% of Cluster 3.

Dataset	Cluster	Number of Samples	Proportion of Dataset in each Cluster (%)	Proportion of Cluster in each Dataset (%)
COLLINS	1	158	10.2	2.4
COLLINS	2	317	20.4	5.2
COLLINS	3	907	58.4	33.9
COLLINS	4	170	11.0	3.1
ELDO	1	2	0.3	0
ELDO	2	101	16.0	1.7
ELDO	3	494	78.3	18.5
ELDO	4	34	5.4	0.6
KERN	1	20	5.3	0.3
KERN	2	32	8.4	0.5
KERN	3	259	68.3	9.7
KERN	4	68	17.9	1.2
KIR	1	6440	35.7	96.9
KIR	2	5551	30.8	91.8
KIR	3	861	4.8	32.2
KIR	4	5166	28.7	94.2
STAN	1	25	9.3	0.4
STAN	2	43	16.0	0.7
STAN	3	156	58.0	5.8
STAN	4	45	16.7	0.8

8. Figures

Figure 1. Relative SDI (%) by assigned SDI_{MAX} type (Pine, Xeric, and Mesic Mixed-Conifer). Below each half-eye plot, dots show mean values, while thick lines show interquartile range (25^{th}-75^{th} percentiles), and thin lines show central 95% of the data (2.5 to 97.5 percentiles). Vertical dotted lines show 25%, 35%, and 60% relative SDI from left to right, respectively.

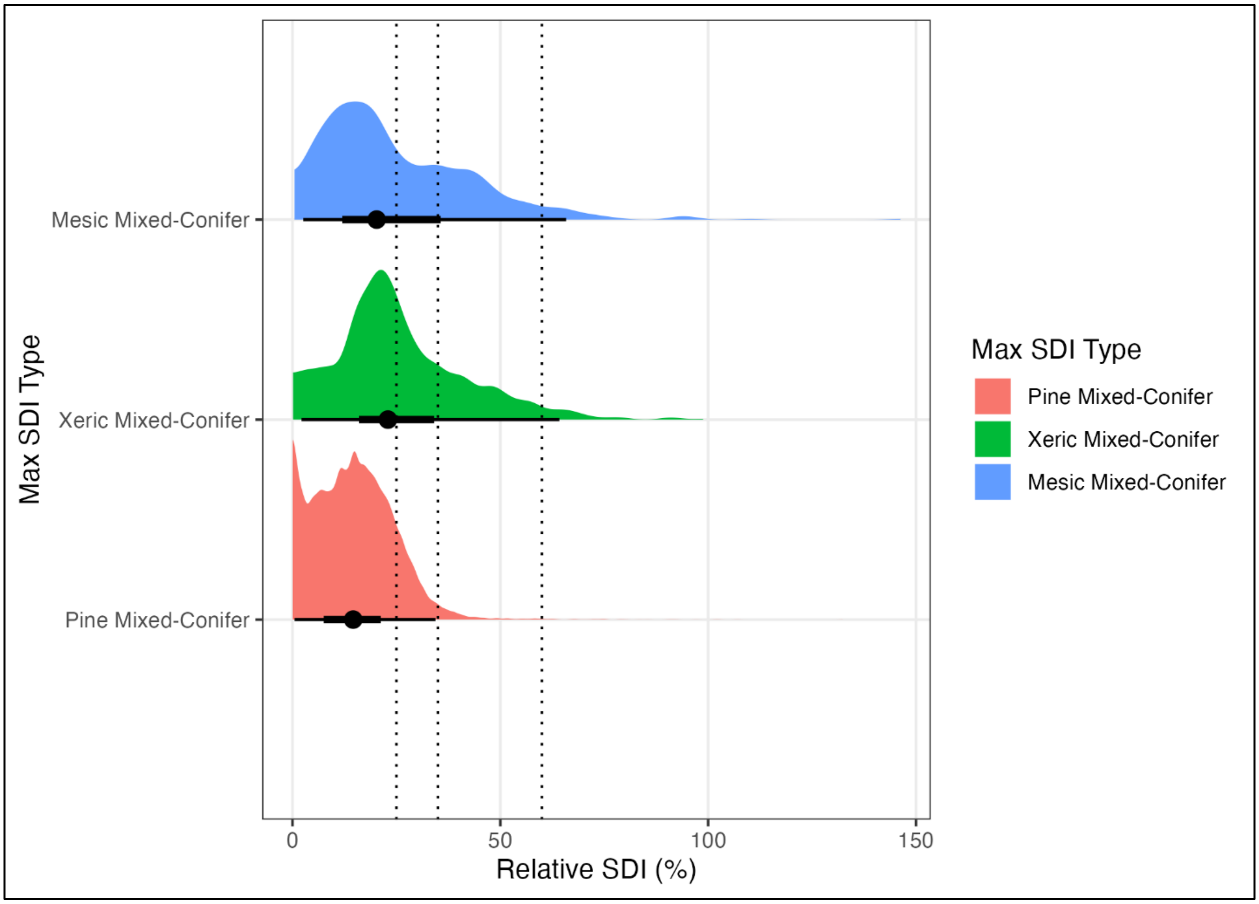


Figure 2. Relative SDI (%) by dataset. Below each half-eye plot, dots show mean values, while thick lines show interquartile range (25th-75th percentiles), and thin lines show central 95% of the data (2.5 to 97.5 percentiles). Vertical dotted lines show 25%, 35%, and 60% relative SDI from left to right, respectively.

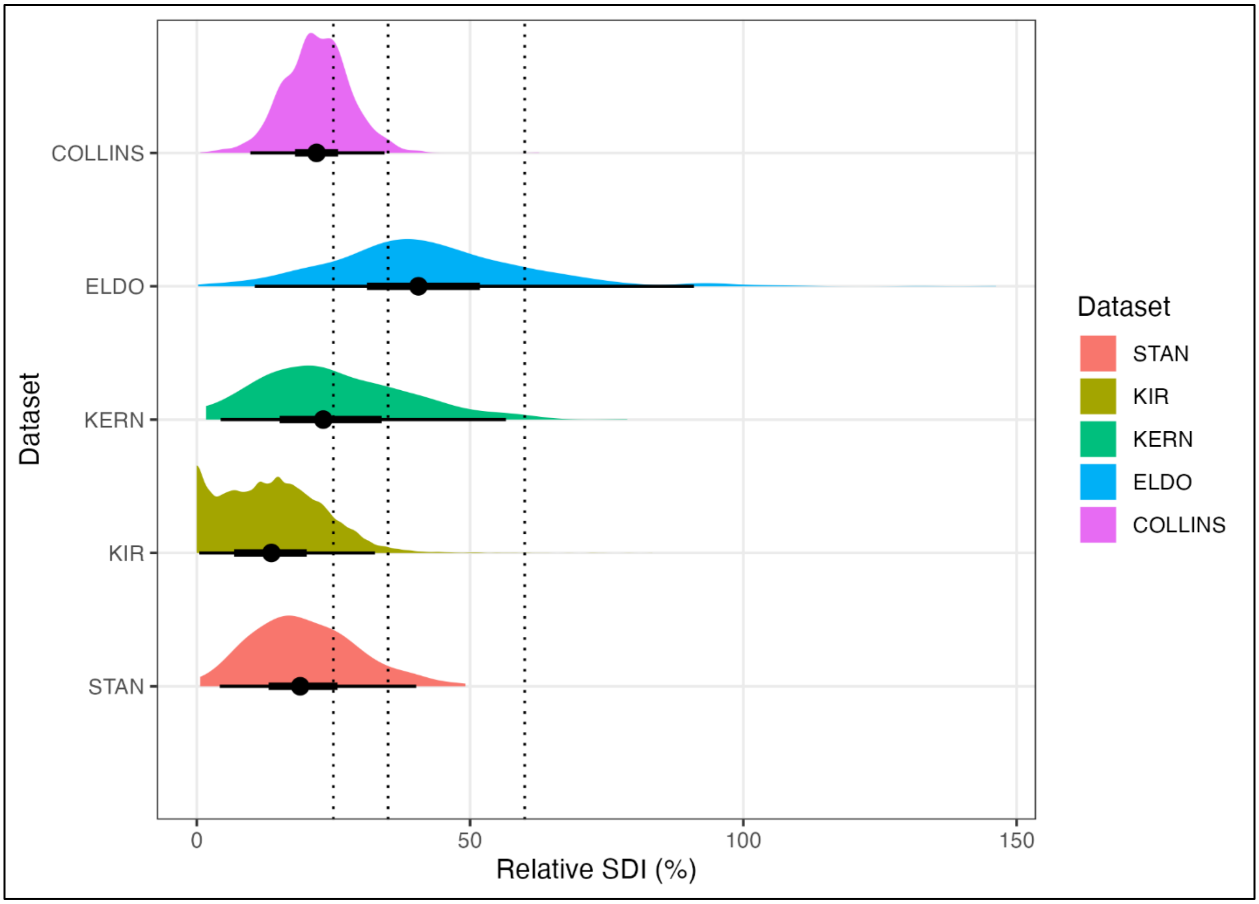


Figure 3. Diameter distributions for datasets with tree-level data (KERN, STAN, KIR). Diameters are binned into 5.1cm classes starting at 15.2cm (bars centered on class means). Dotted red line shows 30.5cm upper diameter limit applied in analyses. Due to display choices, half of the data in the 30.5cm class appears to be below the 30.5cm limit (bar at 30.5 bisected by dotted red line) but only classes entirely below 30.5 were excluded.

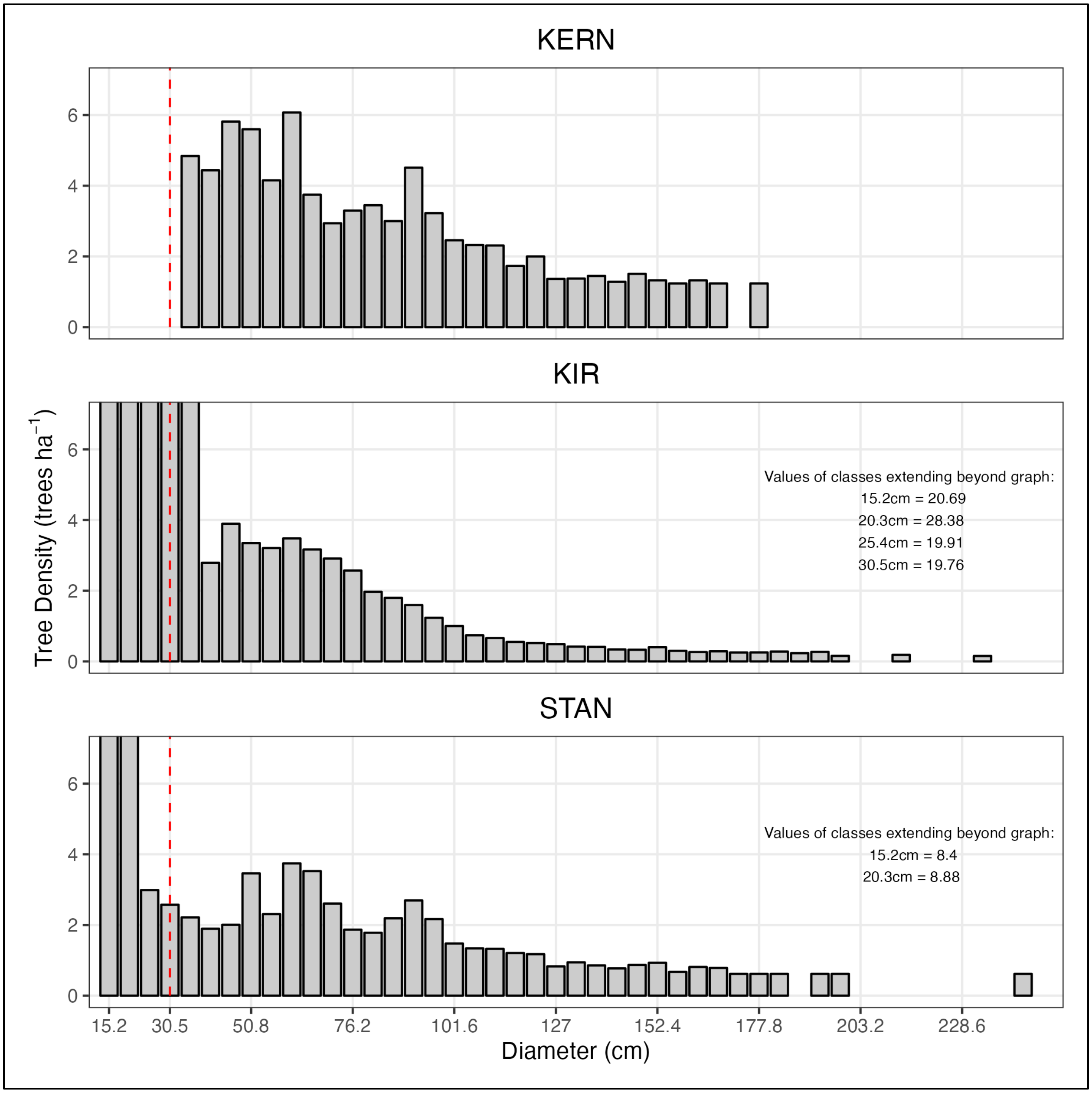
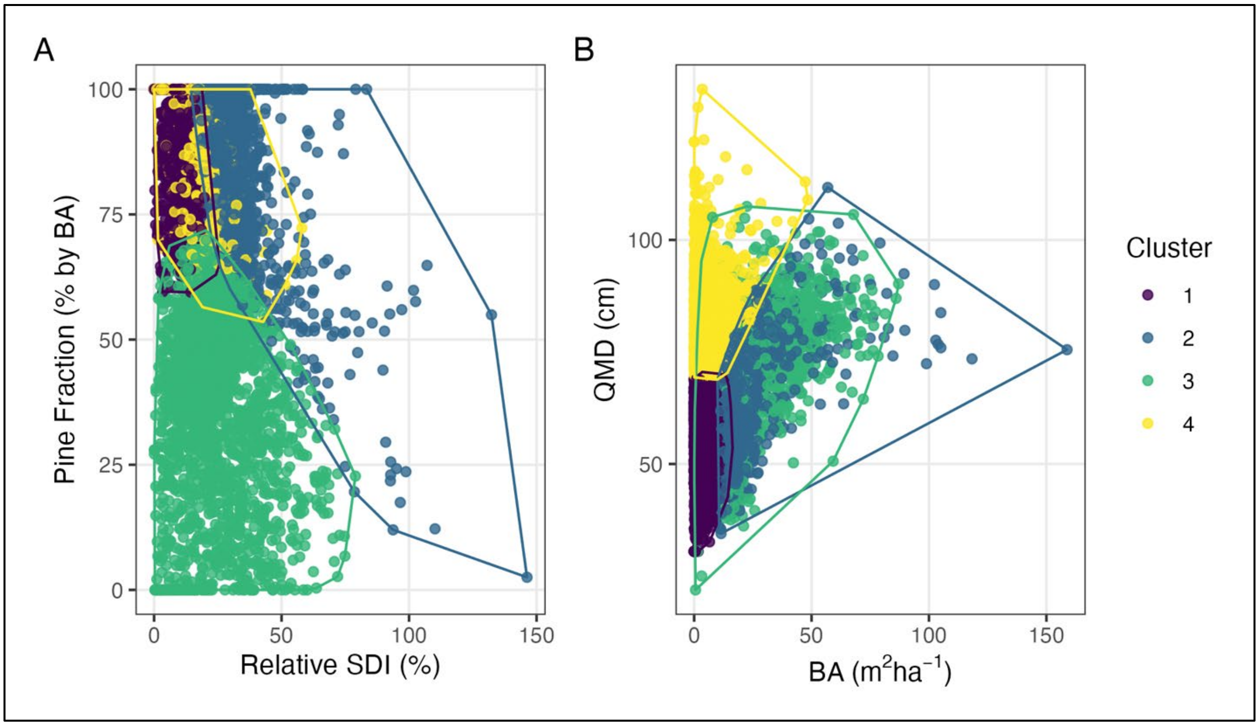


Figure 4. Distributions of key structure and composition variables for clusters identified in K-means analysis. Individual dots represent individual samples, while lines show the convex hull bounding each cluster. Clusters may be descriptively characterized by pine fraction (A), level of competition and stocking (A, B), or QMD (B).



9. References

Anderson, K. 2005. Tending the wild: Native American knowledge and the management of California's natural resources. Univ of California Press.

Barth, M. A. F., A. J. Larson, and J. A. Lutz. 2015. A forest reconstruction model to assess changes to Sierra Nevada mixed-conifer forest during the fire suppression era. Forest Ecology and Management 354:104–118.

Blomdahl, E. M., C. A. Kolden, A. J. H. Meddens, and J. A. Lutz. 2019. The importance of small fire refugia in the central Sierra Nevada, California, USA. Forest Ecology and Management 432:1041–1052.

Bohlman, G. N., H. D. Safford, and C. N. Skinner. 2021. Natural range of variation for yellow pine and mixed-conifer forests in northwestern California and southwestern Oregon. Gen. Tech. Rep. PSW-GTR-273. Albany, CA: U.S. Department of Agriculture, Forest Service, Pacific Southwest Research Station. 146 p. 273.

Cailleret, M., V. Dakos, S. Jansen, E. M. R. Robert, T. Aakala, M. M. Amoroso, J. A. Antos, C. Bigler, H. Bugmann, M. Caccianaga, J.-J. Camarero, P. Cherubini, M. R. Coyea, K. Čufar, A. J. Das, H. Davi, G. Gea-Izquierdo, S. Gillner, L. J. Haavik, H. Hartmann, A.-M. Hereş, K. R. Hultine, P. Janda, J. M. Kane, V. I. Kharuk, T. Kitzberger, T. Klein, T. Levanic, J.-C. Linares, F. Lombardi, H. Mäkinen, I. Mészáros, J. M. Metsaranta, W. Oberhuber, A. Papadopoulos, A. M. Petritan, B. Rohner, G. Sangüesa-Barreda, J. M. Smith, A. B. Stan, D. B. Stojanovic, M.-L. Suarez, M. Svoboda, V. Trotsiuk, R. Villalba, A. R. Westwood, P. H. Wyckoff, and J. Martínez-Vilalta. 2019. Early-warning signals of individual tree mortality based on annual radial growth. Frontiers in Plant Science 9.

Caprio, A., and T. Swetnam. 1995. Historic fire regimes along an elevational gradient on the west slope of the Sierra Nevada, California. Pages. 173–179. Page JK Brown, RW Mutch, CW Spoon and RH Wakimoto [Tech. Coor.], Symposium on fire in wilderness and park management: proceedings. USDA For. Ser. Gen. Tech. Rep. INT-GTR-320.

Carlson, G. T. 1979. Soil resource inventory: Winema National Forest. Department of Agriculture, Forest Service, Pacific Northwest Region.

Chivhenge, E., D. G. Ray, A. R. Weiskittel, C. W. Woodall, and A. W. D'Amato. 2024. Evaluating the Development and Application of Stand Density Index for the Management of Complex and Adaptive Forests. Current Forestry Reports.

Collins, B. M., A. Bernal, R. A. York, J. T. Stevens, A. Juska, and S. L. Stephens. 2021. Mixed-conifer forest reference conditions for privately owned timberland in the southern Cascade Range. Ecological Applications 31:e02400.

Collins, B. M., J. M. Lydersen, R. G. Everett, D. L. Fry, and S. L. Stephens. 2015. Novel characterization of landscape-level variability in historical vegetation structure. Ecological Applications 25:1167–1174.

Das, A. J., N. L. Stephenson, and K. P. Davis. 2016. Why do trees die? Characterizing the drivers of background tree mortality. Ecology 97:2616–2627.

Drew, T. J., and J. W. Flewelling. 1979. Stand Density Management: an Alternative Approach and Its Application to Douglas-fir Plantations. Forest Science 25:518–532.

Ducey, M. J. 2009. The Ratio of Additive and Traditional Stand Density Indices. Western Journal of Applied Forestry 24:5–10.

Egan, J., R. Staff, and J. Blackford. 2015. A high-precision age estimate of the Holocene Plinian eruption of Mount Mazama, Oregon, USA. The Holocene 25:1054–1067.

Flint, L. E., A. L. Flint, J. H. Thorne, and R. Boynton. 2013. Fine-scale hydrologic modeling for regional landscape applications: the California Basin Characterization Model development and performance. Ecological Processes 2:25.

Franklin, J. F., H. H. Shugart, and M. E. Harmon. 1987. Tree Death as an Ecological Process. BioScience 37:550–556.

Garnier, Simon, Ross, Noam, Rudis, Robert, Camargo, A. Pedro, Sciaini, Marco, Scherer, and Cédric. 2024. viridis(Lite) - Colorblind-Friendly Color Maps for R.

Gotelli, N. J., and A. M. Ellison. 2013. A primer of ecological statistics. Second edition. Sinauer Associates, Inc., Publishers, Sunderland, Massachusetts.

Greenler, S. M., F. K. Lake, W. Tripp, K. McCovey, A. Tripp, L. G. Hillman, C. J. Dunn, S. J. Prichard, P. F. Hessburg, W. Harling, and J. D. Bailey. 2024. Blending Indigenous and western science: Quantifying cultural burning impacts in Karuk Aboriginal Territory. Ecological Applications n/a:e2973.

Hagmann, R. K., J. F. Franklin, and K. N. Johnson. 2013. Historical structure and composition of ponderosa pine and mixed-conifer forests in south-central Oregon. Forest Ecology and Management 304:492–504.

Hagmann, R. K., P. F. Hessburg, S. J. Prichard, N. A. Povak, P. M. Brown, P. Z. Fulé, R. E. Keane, E. E. Knapp, J. M. Lydersen, K. L. Metlen, M. J. Reilly, A. J. Sánchez Meador, S. L. Stephens, J. T. Stevens, A. H. Taylor, L. L. Yocom, M. A. Battaglia, D. J. Churchill, L. D. Daniels, D. A. Falk, P. Henson, J. D. Johnston, M. A. Krawchuk, C. R. Levine, G. W. Meigs, A. G. Merschel, M. P. North, H. D. Safford, T. W. Swetnam, and A. E. M. Waltz. 2021. Evidence for widespread changes in the structure, composition, and fire regimes of western North American forests. Ecological Applications 31:e02431.

Hagmann, R. K., D. L. Johnson, and K. N. Johnson. 2017. Historical and current forest conditions in the range of the Northern Spotted Owl in south central Oregon, USA. Forest Ecology and Management 389:374–385.

Hagmann, R. K., A. G. Merschel, and M. J. Reilly. 2019. Historical patterns of fire severity and forest structure and composition in a landscape structured by frequent large fires: Pumice Plateau ecoregion, Oregon, USA. Landscape Ecology 34:551–568.

Hagmann, R. K., J. T. Stevens, J. M. Lydersen, B. M. Collins, J. J. Battles, P. F. Hessburg, C. R. Levine, A. G. Merschel, S. L. Stephens, A. H. Taylor, J. F. Franklin, D. L. Johnson, and K. N. Johnson. 2018. Improving the use of early timber inventories in reconstructing historical dry forests and fire in the western United States: Comment. Ecosphere 9:e02232.

Kilgore, B. M., and D. Taylor. 1979. Fire History of a Sequoia-Mixed Conifer Forest. Ecology 60:129–142.

Kimsey, M. J., T. M. Shaw, and M. D. Coleman. 2019. Site sensitive maximum stand density index models for mixed conifer stands across the Inland Northwest, USA. Forest Ecology and Management 433:396–404.

Knapp, E. E., C. N. Skinner, M. P. North, and B. L. Estes. 2013. Long-term overstory and understory change following logging and fire exclusion in a Sierra Nevada mixed-conifer forest. Forest Ecology and Management 310:903–914.

Knight, C. A., C. V. Cogbill, M. D. Potts, J. A. Wanket, and J. J. Battles. 2020. Settlement-era forest structure and composition in the Klamath Mountains: reconstructing a historical baseline. Ecosphere 11:e03250.

LANDFIRE, E. R. O. and S. C. (EROS). 2022, January 31. LANDFIRE 2020 Elevation (Elev) CONUS. data.raster digital data, LANDFIRE, Earth Resources Observation and Science Center (EROS), U.S. Geological Survey. https://landfire.gov/metadata/lf2020/CONUS/LC20_Elev_220.html#7.

Long, J. N. 1985. A Practical Approach to Density Management. The Forestry Chronicle 61:23–27.

Long, J. N., and T. W. Daniel. 1990. Assessment of growing stock in uneven-aged stands. Western Journal of Applied Forestry 5:93–96.

Long, J. N., and J. D. Shaw. 2005. A density management diagram for even-aged ponderosa pine stands. Western Journal of Applied Forestry 20:205–215.

Long, J. W., F. K. Lake, and R. W. Goode. 2021. The importance of Indigenous cultural burning in forested regions of the Pacific West, USA. Forest Ecology and Management 500:119597.

Ma, S., A. Concilio, B. Oakley, M. North, and J. Chen. 2010. Spatial variability in microclimate in a mixed-conifer forest before and after thinning and burning treatments. Forest Ecology and Management 259:904–915.

Maxwell, R. S., A. H. Taylor, C. N. Skinner, H. D. Safford, R. E. Isaacs, C. Airey, and A. B. Young. 2014. Landscape-scale modeling of reference period forest conditions and fire behavior on heavily logged lands. Ecosphere 5:art32.

McIntyre, P. J., J. H. Thorne, C. R. Dolanc, A. L. Flint, L. E. Flint, M. Kelly, and D. D. Ackerly. 2015. Twentieth-century shifts in forest structure in California: Denser forests, smaller trees, and increased dominance of oaks. Proceedings of the National Academy of Sciences 112:1458–1463.

Merschel, A. G., E. K. Heyerdahl, T. A. Spies, and R. A. Loehman. 2018. Influence of landscape structure, topography, and forest type on spatial variation in historical fire regimes, Central Oregon, USA. Landscape Ecology 33:1195–1209.

Meyer, M., and M. North. 2019. Natural range of variation of red fir and subalpine forests in the Sierra Nevada Bioregion. Gen Tech. Rep. PSW-GTR-263. Albany, CA: U.S. Department of Agriculture, Forest Service, Pacific Southwest Research Station 263.

NARA, 1914-22. Timber Inventory Tally Sheets. Records of the Bureau of Indian Affairs, Record Group 75. National Archives and Records Building, Seattle.

Nepal, S., B. N. I. Eskelson, M. W. Ritchie, and S. E. Gergel. 2023. Spatial patterns of vigor by stand density across species groups and its drivers in a pre-harvest ponderosa pine-dominated landscape in northern California. Forest Ecology and Management 534:120867.

North, M., J. Innes, and H. Zald. 2007. Comparison of thinning and prescribed fire restoration treatments to Sierran mixed-conifer historic conditions. Canadian Journal of Forest Research 37:331–342.

North, M. P., R. E. Tompkins, A. A. Bernal, B. M. Collins, S. L. Stephens, and R. A. York. 2022. Operational resilience in western US frequent-fire forests. Forest Ecology and Management 507:120004.

North, M., P. Stine, K. O'Hara, W. Zielinski, and S. Stephens. 2009. An ecosystem management strategy for Sierran mixed-conifer forests. Page PSW-GTR-220. U.S. Department of Agriculture, Forest Service, Pacific Southwest Research Station, Albany, CA.

O'Geen, A. T., R. A. Dahlgren, and D. Sanchez-Mata. 2007. California soils and examples of ultramafic vegetation. Terrestrial vegetation of California:456–501.

Oksanen, J., G. L. Simpson, F. G. Blanchet, R. Kindt, P. Legendre, P. R. Minchin, R. B. O'Hara, P. Solymos, M. H. H. Stevens, E. Szoecs, H. Wagner, M. Barbour, M. Bedward, B. Bolker, D. Borcard, G. Carvalho, M. Chirico, M. D. Caceres, S. Durand, H. B. A. Evangelista, R. FitzJohn, M. Friendly, B. Furneaux, G. Hannigan, M. O. Hill, L. Lahti, D. McGlinn, M.-H. Ouellette, E. R. Cunha, T. Smith, A. Stier, C. J. F. T. Braak, and J. Weedon. 2022, October 11. vegan: Community Ecology Package.

Palik, B. J., A. W. D'Amato, J. F. Franklin, and K. N. Johnson. 2020. Ecological silviculture: Foundations and applications. Waveland Press.

Pedersen, T. L. 2024. patchwork: The Composer of Plots.

Posit Team. 2022. RStudio: Integrated Development Environment for R. Posit Software, PBC, Boston, MA.

PRISM Climate Group, Oregon State University, http://prism.oregonstate.edu, created 6 Dec 2006.

R Core Team. 2022. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Reineke, L. H. 1933. Perfecting a stand-density index for even-aged forests. Journal of Agricultural Research 46:627–638.

Ritchie, M. W. 2016. Multi-scale reference conditions in an interior pine-dominated landscape in northeastern California. Forest Ecology and Management 378:233–243.

Safford, H. D., and J. T. Stevens. 2017. Natural range of variation for yellow pine and mixed-conifer forests in the Sierra Nevada, southern Cascades, and Modoc and Inyo National Forests, California, USA. Gen. Tech. Rep. PSW-GTR-256. Albany, CA: U.S. Department of Agriculture, Forest Service, Pacific Southwest Research Station. 229 p. 256.

Safford, H., M. North, and M. Meyer. 2012. Climate change and the relevance of historical forest conditions. Managing Sierra Nevada forests. Gen. Tech. Rep. PSW-GTR-237. Albany, CA: US Department of Agriculture, Forest Service, Pacific Southwest Research Station:23–46.

Scholl, A. E., and A. H. Taylor. 2010. Fire regimes, forest change, and self-organization in an old-growth mixed-conifer forest, Yosemite National Park, USA. Ecological Applications 20:362–380.

Sensenig, T., J. D. Bailey, and J. C. Tappeiner. 2013. Stand development, fire and growth of old-growth and young forests in southwestern Oregon, USA. Forest Ecology and Management 291:96–109.

Shaw, J. D. 2000. Application of Stand Density Index to Irregularly Structured Stands. Western Journal of Applied Forestry 15:40–42.

Skinner, C., A. Taylor, and J. Agee. 2018. Klamath Mountains bioregion. In 'Fire in California's Ecosystems', 2nd edn.(Eds JW van Wagtendonk, NG Sugihara, SL Stephens, AE Thode, KE Shaffer, J Fites) pp. 170–194. University of California Press: Berkeley, CA, USA.

Stephens, S. L., and B. M. Collins. 2004. Fire regimes of mixed conifer forests in the north-central Sierra Nevada at multiple spatial scales. Northwest Science 78:12–23.

Stephens, S. L., L. Hall, C. W. Stephens, A. A. Bernal, and B. M. Collins. 2023. Degradation and restoration of Indigenous California black oak (Quercus kelloggii) stands in the northern Sierra Nevada. Fire Ecology 19:12.

Stephens, S. L., J. M. Lydersen, B. M. Collins, D. L. Fry, and M. D. Meyer. 2015. Historical and current landscape-scale ponderosa pine and mixed conifer forest structure in the Southern Sierra Nevada. Ecosphere 6:1–63.

Stephens, S. L., J. T. Stevens, B. M. Collins, R. A. York, and J. M. Lydersen. 2018. Historical and modern landscape forest structure in fir (Abies)-dominated mixed conifer forests in the northern Sierra Nevada, USA. Fire Ecology 14:7.

Stephenson, N. 1998. Actual evapotranspiration and deficit: biologically meaningful correlates of vegetation distribution across spatial scales. Journal of Biogeography 25:855–870.

SUGIHARA, N. G., J. W. V. WAGTENDONK, and J. A. FITES-KAUFMAN. 2018. Fire as an Ecological Process. Pages 57–70 Fire in California's Ecosystems. Second edition. University of California Press.

Taylor, A. H. 2004. Identifying Forest Reference Conditions on Early Cut-Over Lands, Lake Tahoe Basin, Usa. Ecological Applications 14:1903–1920.

Taylor, A. H., V. Trouet, C. N. Skinner, and S. Stephens. 2016. Socioecological transitions trigger fire regime shifts and modulate fire–climate interactions in the Sierra Nevada, USA, 1600–2015 CE. Proceedings of the National Academy of Sciences 113:13684–13689.

Taylor, A. H., A. M. Vandervlugt, R. S. Maxwell, R. M. Beaty, C. Airey, and C. N. Skinner. 2014. Changes in forest structure, fuels and potential fire behaviour since 1873 in the Lake Tahoe Basin, USA. Applied Vegetation Science 17:17–31.

Van de Water, K., and M. North. 2011. Stand structure, fuel loads, and fire behavior in riparian and upland forests, Sierra Nevada Mountains, USA; a comparison of current and reconstructed conditions. Forest Ecology and Management 262:215–228.

Wickham, H., M. Averick, J. Bryan, W. Chang, L. D. McGowan, R. François, G. Grolemund, A. Hayes, L. Henry, J. Hester, M. Kuhn, T. L. Pedersen, E. Miller, S. M. Bache, K. Müller, J. Ooms, D. Robinson, D. P. Seidel, V. Spinu, K. Takahashi, D. Vaughan, C. Wilke, K. Woo, and H. Yutani. 2019. Welcome to the tidyverse. Journal of Open Source Software 4:1686.

Wieslander, A. E. 1935. A Vegetation Type Map of California. Madroño 3:140–144.

Woodall, C. W., A. W. D’Amato, J. B. Bradford, and A. O. Finley. 2011. Effects of Stand and Inter-Specific Stocking on Maximizing Standing Tree Carbon Stocks in the Eastern United States. Forest Science 57:365–378.

Woodall, C. W., and A. R. Weiskittel. 2021. Relative density of United States forests has shifted to higher levels over last two decades with important implications for future dynamics. Scientific Reports 11:18848.

Ziegler, J. P., C. M. Hoffman, B. M. Collins, E. E. Knapp, and W. (Ruddy) Mell. 2021. Pyric tree spatial patterning interactions in historical and contemporary mixed conifer forests, California, USA. Ecology and Evolution 11:820–834.