Comparing methods of reconstructing fire history using fire scars in a southwestern United States ponderosa pine forest

Megan L. Van Horne and Peter Z. Fulé

Abstract: Fire scars have been used to understand the historical role of fire in ponderosa pine (*Pinus ponderosa* Dougl. ex P. & C. Laws.) ecosystems, but sampling methods and interpretation of results have been criticized for being statistically invalid and biased and for leading to exaggerated estimates of fire frequency. We compared "targeted" sampling, random sampling, and grid-based sampling to a census of all 1479 fire-scarred trees in a 1 km^2 study site in northern Arizona. Of these trees, 1246 were sufficiently intact to collect cross-sections; of these, 648 had fire scars that could be cross-dated to the year of occurrence in the 200-year analysis period. Given a sufficient sample size (approximately $n \geq 50$), we concluded that all tested sampling methods resulted in accurate estimates of the census fire frequency, with mean fire intervals within 1 year of the census mean. We also assessed three analytical techniques: (1) fire intervals from individual trees, (2) the interval between the tree origin and the first scar, and (3) proportional filtering. "Bracketing" fire regime statistics to account for purported uncertainty associated with targeted sampling was not useful. Quantifying differences in sampling approaches cannot resolve all the limitations of fire-scar methods, but does strengthen interpretation of these data.

Résumé : Les cicatrices de feu ont été utilisées pour comprendre le rôle que le feu avait jadis dans les écosystèmes de pin ponderosa (*Pinus ponderosa* Dougl. ex P. & C. Laws.) mais les méthodes d'échantillonnage et l'interprétation des résultats ont été remises en cause et considérées comme invalides du point de vue statistique, biaisées et menant à des estimations exagérées de la fréquence des feux. Nous avons comparé l'échantillonnage ciblé, l'échantillonnage aléatoire et l'échantillonnage en grille à un recensement des 1479 arbres portant des cicatrices de feu sur un site de 1 km^2 dans le nord de l'Arizona. De ces arbres, 1246 étaient suffisamment intacts pour prélever des disques parmi lesquels 648 portaient des cicatrices de feu dont l'origine pouvait être datée au cours de la période d'analyse de 200 ans. Étant donné que la taille de l'échantillon était suffisante ($n \geq 50$ approximativement), nous avons conclu que toutes les méthodes d'échantillonnage testées produisaient des estimations justes de la fréquence des feux recensés avec des intervalles moyens entre les feux dont l'écart à la moyenne du recensement était inférieure à un an. Nous avons également évalué trois techniques d'analyses : (1) les intervalles entre les feux à partir d'arbres individuels, (2) l'intervalle entre l'origine des arbres et la première cicatrice de feu et (3) le filtrage proportionnel. L'utilisation de la méthode d'encadrement des statistiques de régime de feux pour tenir compte de la prétendue incertitude associée à l'échantillonnage ciblé n'était pas utile. Le fait de quantifier les différences entre les modes d'échantillonnage ne peut résoudre toutes les déficiences des méthodes utilisées pour les cicatrices de feu mais renforce l'interprétation de ces données.

[Traduit par la Rédaction]

Introduction

Fire scars provide a temporally precise way to reconstruct fire history, but uncertainties are inherent when estimating fire frequency and spatial patterning of fires for three main reasons: (1) fire scars are not consistently recorded on individual scarred trees, so they are an incomplete point source of data (Dieterich and Swetnam 1984), (2) more recent fire events may have consumed remnant fire records, and (3) error is introduced by the process of sampling the population of fire scars (Fall 1998). Few researchers have attempted to quantify the extent of uncertainty in fire-scar studies, though most acknowledge that problems exist (e.g., Swetnam and Baisan 1996; Fall 1998). Baker and Ehle (2001) suggested "bracketing" mean fire intervals (MFI) with correction factors to compensate for the perceived uncertainties in 18 previously published studies in ponderosa pine (*Pinus ponderosa* Dougl. ex P. & C. Laws.) forests that reported MFI values of 5–21 years. When their bracketing methods were applied, Baker and Ehle (2001) calculated the MFI to be 22–308 years. In contrast, Fall (1998) argued that current methods are biased in the opposite direction, towards underrepresenting fire occurrence because many unscarred trees may have actually burned but failed to scar. To explore several factors

Received 31 May 2005. Accepted 26 November 2005. Published on the NRC Research Press Web site at http://cjfr.nrc.ca on 30 March 2006.

M.L. Van Horne and P.Z. Fulé.[1,2] Ecological Restoration Institute and School of Forestry, Northern Arizona University, P.O. Box 15018, Flagstaff, AZ 86011, USA.

[1]Corresponding author (e-mail: Pete.Fule@nau.edu).
[2]Present address: Southwest Forest Science Complex, 82 Huffer Lane, room 116, Flagstaff, AZ 86011, USA.

 doi:10.1139/X05-289

that have been raised as "uncertainties" in reconstructing ponderosa pine fire history from fire scars, we initiated a complete census of a study site.

Sampling of fire-scarred trees is often "targeted" toward trees showing multiple scars and long records of fire to compile a complete inventory of fire years in that area (Arno and Sneck 1977; McBride 1983; Agee 1993; Swetnam and Baisan 1996; Fulé et al. 1997, Swetnam and Baisan 2003; Grissino-Mayer et al. 2004) Targeted samples can be collected from throughout a study area or near systematic grid points (Arno et al. 1995; Heyerdahl 1997). Targeting has been criticized for being statistically invalid because it is not a random sample from a well-defined population, and therefore it can lead to estimates of fire frequency where neither the accuracy nor the precision is known (Johnson and Gutsell 1994; Baker and Ehle 2003). Swetnam and Baisan (1996) argued that random sampling would not result in a complete or unbiased record of fire in frequent surface fire regimes unless very large numbers of trees were sampled. They supported the targeting method based on the argument that trees are a natural archive of historical data and not consistently reliable recorders of fire, so they should not be treated "as if they all belong to the same statistical population" (Swetnam and Baisan 1996, p. 13).

Composite fire intervals are used to capture the complete record of fire dates in an area. A composite (Dieterich 1980) fire interval analysis typically results in a much shorter MFI compared with that of an individual tree, or point fire interval. While Baker and Ehle (2001) suggested that the composite MFI overestimates fire occurrence and is not area explicit, the point MFI likely underestimates fire occurrence because of unrecorded fires. One approach to resolving this problem is filtering the composite by including only fire dates that occur on greater than a determined percentage of trees (Grissino-Mayer 1995). Ten percent and 25% filtered composites have been used to represent the fires that are likely to be progressively more widespread within a study site (Swetnam and Baisan 1996; Baker and Ehle 2001).

The period between tree germination and the first fire scar, called the origin-to-scar (OS) interval (Baker 1989), is another potential source of uncertainty. Baker and Ehle (2001, 2003) argued that for a ponderosa pine tree to survive, it must have experienced a fire-free interval of at least 50 years, and therefore this fire-free interval must be included in the estimate of MFI. Stephens et al. (2003) countered their argument by maintaining that it is impossible to know the true fire-free interval because many trees survive fires without ever scarring. Furthermore, because older fire scars or even the tree pith may be burned away by subsequent fires, the OS interval often cannot be measured (Stephens et al. 2003).

Johnson and Gutsell (1994) asserted that the use of time-since-fire maps to calculate fire rotation (FR, the time required to burn over an area equal to the study area) is the only statistically valid method of reconstructing fire events because it accounts for spatial and temporal variability. Baker and Ehle (2001) argued that the fire interval should be interpreted as being equivalent to the FR. An underlying assumption is that point data (fire-scarred trees) can represent a particular area burned. If a series of fire scars formed in the same year indicate that a large area burned, FR may be relatively short, as would MFI. But if the scarred trees represent a number of small, localized fires, as argued by Minnich et al. (2000), and if unscarred trees or those that had not yet received their first fire scar were considered evidence of the absence of fire (Baker and Ehle 2001), then results obtained with the MFI and FR methods would diverge.

The preceding concerns raised in the literature regarding fire-scar data led us to design a study to census and map an entire population of fire-scarred ponderosa pine trees. In this case study we ask the following questions:

(1) Are targeting and other methods of sampling (random, grid based) accurate when compared with census data?
(2) How do sample size, area sampled, and filtering affect MFI estimates?
(3) Should OS intervals, "bracketed" intervals, and point MFIs be included in fire history?
(4) Previous fire ecology studies in this region concluded that surface fires were frequent and led to recommendations for thinning and burning for forest restoration. Are our results consistent with these interpretations?

Methods

Study site

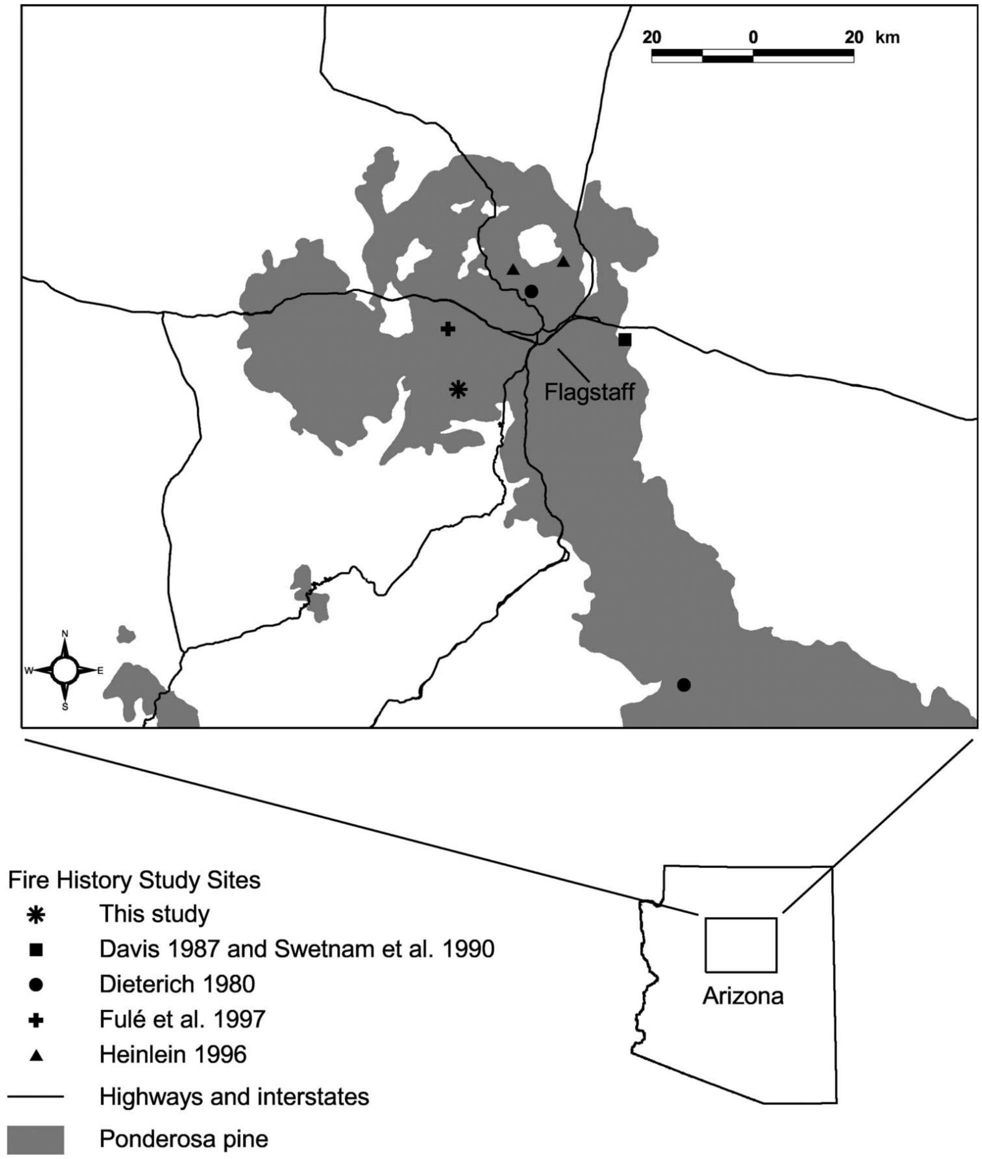
We censused fire-scarred trees on 1 km^2 in Northern Arizona University's Centennial Forest approximately 20 km southwest of Flagstaff, Arizona (35°05′N, 111°50′W) (Fig. 1). Given the extraordinary effort required to census fire scars, this is necessarily a case study. However, the landscape features, stand conditions, and land-use history of the site are representative of much of the ponderosa pine forest type in Arizona and New Mexico. Located north of the Mogollon Rim at 2200 m elevation, this site was selected for its relatively low variation in vegetation and topography and the lack of natural barriers to fire spread. The soils are predominately sandy loams and loams (Abella 2005) with basalt cinders and limestone parent material (Miller et al. 1995). Average annual precipitation in Flagstaff is approximately 54 cm (1950–2004) with most of the precipitation falling in late winter and late summer (Western Regional Climate Center 2004). Ponderosa pine is the dominant species with occasional Gambel oak (*Quercus gambelii* Nutt.) in the understory. Basal area averages 19.1 m^2/ha (range 0 to 45.9 m^2/ha). Slopes range from 0% to 10%.

Timber extraction began in the site and surrounding areas during railroad construction in the 1880s, and at about the same time overgrazing led to widespread exclusion of fire (Dieterich 1980; Fulé et al. 1997) Sporadic logging activities and fire suppression continued throughout the 20th century. The most recent activity in the site was timber harvesting and burning piles of logging slash in the 1980s and 1990s.

Field methods

The targeted sample was collected first. We systematically walked through the site examining each fire-scarred tree we encountered for number of visible scars and soundness. Forty trees with multiple scars and long records of fire, distributed throughout the sampling unit, were selected to compose the targeted sample. After completing the targeted sample collection, a grid consisting of twenty-five 4 ha cells was overlaid on the 100 ha study site. The grid was flagged to

Fig. 1. Fire history study sites, including the present study, in northern Arizona.



delineate boundaries for organized collection and analysis of the census. The entire population of trees with visible fire-scar evidence constituted the census. Cross-sections from all remaining fire-scarred trees (those not collected in the targeted sample) were collected in the census. Several specimens were too deteriorated to collect, so the fire-scarred tree was documented and mapped.

We recognized that buried scars may have existed on the site, but we did not deliberately sample into intact living trees for them. However, virtually all the pre 20th century trees sampled were old cut stumps, so in some cases scars that would not have been visible in a standing tree were evident on the stump. Scarred trees included living trees, snags, logs, and stumps. We collected partial cross-sections from live and standing dead trees, a standard technique that can be done without killing live trees (Arno and Sneck 1977; Heyerdahl and McKay 2001). A 5 cm thick cross-section was extracted from the region of the tree that appeared to have the most complete fire record using a chainsaw. Complete cross-sections were cut from stumps and logs. In cases where multiple sides or heights on the fire-scarred tree appeared to have recorded different fires, multiple cross-sections from that tree were extracted. The only living trees with fire scars were young trees clustered around burned slash piles, and they appeared to be scarred recently. Because we were primarily interested in the fire regime before Euro-American settlement, we only collected samples from about 70% of these living trees to verify that the fire date was outside our time frame of analysis. However, all these trees were measured and mapped in the field.

Each fire-scarred tree was documented by recording diameter at stump height, number of cross-sections taken, number of pieces per cross-section, number of visible scars on the specimen, aspect of the fire-scarred area(s) on the tree, height of the cross-sections on the bole, and Universal Transverse Mercator (UTM) coordinates from a Garmin® GPS, accurate to within 15 m. The condition of the tree was also recorded as living, snag, stump, or log.

Laboratory methods

All specimens were mounted on plywood and surfaced using an electric belt sander with increasingly finer sandpaper until the cells were clearly visible under magnification. We used a ring-width chronology from the Gus Pearson Natural Area, Arizona (Graybill 1987), supplemented with 20 increment cores from old trees on the site, to build a master ring-width chronology specific to the study site. All specimens were visually cross-dated when possible, using visual recognition of tree-ring patterns supported by skeleton plotted chronologies and lists of marker years (those with narrow rings) (Stokes and Smiley 1968). COFECHA software (Holmes 1983) was used to assist with dating difficult specimens. The rings on the difficult specimens were measured with an Acu-Rite glass scale and encoder with 2 µm precision and Measure J2X software. Dating suggested by COFECHA outputs was checked carefully against the original cross-sections to verify the dating visually, rather than relying solely on correlation coefficients. Many specimens remained undated even after COFECHA was employed. If a specimen had an injury that could not be unquestionably identified as a fire scar, we did not include the date of that injury. If a fire scar could not be dated to an exact year, we did not include the estimated year. Specimens that contained no fire scars or had fire scars where the exact year could not be determined were not cross-dated.

All the targeted specimens and 40% of the entire collection were checked by other dendrochronologists to independently verify the dates of the wood specimen and fire events. Any unresolved discrepancies were considered undateable. We also identified the years in which each tree with dateable fire scars was recording. A tree is considered to be “recording” after the initial injury because the wounded area is susceptible to being rescarred by subsequent fires (Grissino-Mayer 1995).

Data analysis

Each individual fire-scarred tree will be called a specimen (even if more than one cross-section was collected from that tree), and each group of specimens analyzed will be a sample. FHX2 software (Grissino-Mayer 1995) was used to analyze combinations of fire-history data from specimens. The 200-year period from 1682 to 1881 was used in all analyses for consistency unless otherwise stated. The minimum number of trees recording in this period was 39 trees in 1881, a sufficient number of trees with which to conduct this analysis (Falk and Swetnam 2003). The year 1881 is the last year of analysis because it was the last fire year in the study area before grazing and fire suppression interrupted the natural fire regime. Prior to the late 1600s, the results would be confounded by the lack of fire scar data. The year 1682 was chosen to make an even 200-year period of analysis.

We chose the mean fire interval (MFI) as the basic statistic for analysis in this study, following the example of Baker and Ehle (2001), who noted that this statistic was the only one consistently reported throughout the literature and therefore was the best value for cross-study comparisons. This is not intended to imply that MFI is necessarily the most useful measure of central tendency (e.g., see Grissino-Mayer 1999; Falk and Swetnam 2003). For each sample, the MFI for all scars (no filter) and for the 10% and 25% scarred filters was computed in FHX2 (Dieterich 1980; Swetnam and Baisan 1996) Filters only include those 5 years that are recorded by the determined minimum percentage of recording trees and can be used to infer fire size; no filter includes fires of all sizes, whereas a 25% filter only includes the larger fires. These samples are subsets of the census data, so the same specimens may be included in multiple samples. Because of lack of independence and spatial autocorrelation, we did not test for statistically significant differences of the means. Instead, graphical and tabular representations of the means are used to show the effect of sampling methods.

A GIS was used to map the locations of the samples by the UTM coordinates recorded in the field to assist with the spatial interpretations. The original UTM coordinates were used except where the trees falsely appeared to be out of the study area because of GPS error. UTM coordinates of these trees were moved to the site boundary nearest to the original location.

Comparison of sampling methods

The analyses outlined above were applied to the following samples:

(1) Census — The census provided a baseline to test the effect of all other sampling methods. The census fire history was assembled using all scars and represented the most complete possible fire-scar-based fire history of the site.

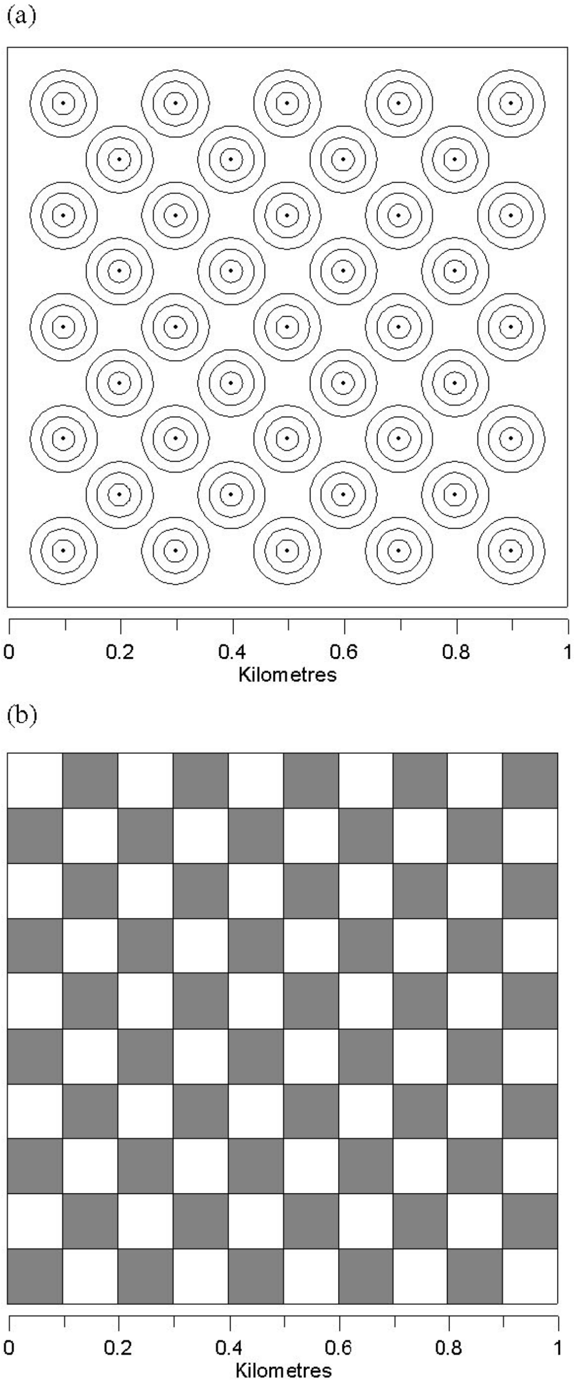
(2) Targeted sample — The targeted sample was analogous to other fire-history studies in the region. The targeted sample consisted of 40 specimens with multiple scars and long records of fire.

(3) Sample size (random samples) — Random samples were used to test the effect of increasing sample size. Eight sample sizes were tested: 10, 20, 30, 40, 50, 60, 70, and 80 specimens. For each sample size analysis, all cross-dated specimens with fires were randomly sampled without replacement from the complete data set using a random number algorithm. Sampling with replacement occurred at the scale of the separate samples, but not within a single sample. For each sample size, the MFIs for 10 random samples were averaged. For samples with ≥40 specimens, we calculated the range and standard deviation of the fire intervals.

(4) Area sampled — The study area was spatially subdivided into 25 grid cells of equal size (200 m × 200 m). This analysis considered six areas of different sizes to test the effect of increasing the size of the study area on MFI. The areas tested were 4 ha (1 cell), 8 ha (2 cells), 16 ha (4 cells), 32 ha (8 cells), 64 ha (16 cells), and 100 ha (25 cells). All dated specimens were included in this analysis. The 100 ha sample is equivalent to the census. The MFI for each area was calculated as an average of the MFIs for each combination of that size. That is, there are four combinations of 16 adjacent cells arranged in a square. Those four combinations were averaged to get the 16-cell MFI; however, the MFI for each of the four combinations is shown graphically.

(5) Grid-based samples — We compared two alternative systematic grid-based sampling approaches. Grid 1 had a spacing of 141 m arranged diagonally over the study area, yielding 41 plots (see Fig. 2*a*). The spacing of

Fig. 2. Maps of sampling grids used to test the effect of grid-based methods of sampling. (*a*) Sampling grid 1 had 41 plots spaced at 141 m. Concentric circles are 20, 40, and 60 m search radii. Mean fire interval (MFI) was compared between samples taken from the three search radii. (*b*) Sampling grid 2 was a checkerboard with one hundred 1 ha cells. MFI of the black cells was compared with MFI of the white cells using the specimen in each cell with the most fire scars.



these plots was determined by the original 25-cell grid, locating one plot in the middle of each cell and one plot at the corner of four adjacent cells. Each sample included one to four specimens with the highest number of observed scars in the field from within three search radii, 20, 40, and 60 m. The samples consisted of different numbers of specimens per plot to approximately mimic the size of the targeted sample (n = 40). Within the 40 m radius, two different samples were assessed: (*i*) one specimen per plot, and (*ii*) two specimens per plot. Grid 2 is a checkerboard with one hundred 1 ha blocks (see Fig. 2*b*). MFIs were calculated using the specimen with the highest number of observed scars in the field, as recorded at the time of specimen collection, from each of the 50 white cells, then from each of the 50 black cells.

Point MFIs and origin-to-first-scar intervals

Point MFIs were calculated for each specimen in the census, then plotted for comparison with the census composites with the three levels of filtering. The OS interval distribution of all 154 specimens with piths was analyzed, and the proportion of OS intervals less than 50 years was reported (Baker and Ehle 2001). We also determined the OS interval for the 47 specimens having piths within the 200-year period of analysis and OS intervals less than the point MFI. We mapped these 47 specimens and assessed whether the OS intervals were truly fire-free, based on the fire dates of their nearest recording neighbors.

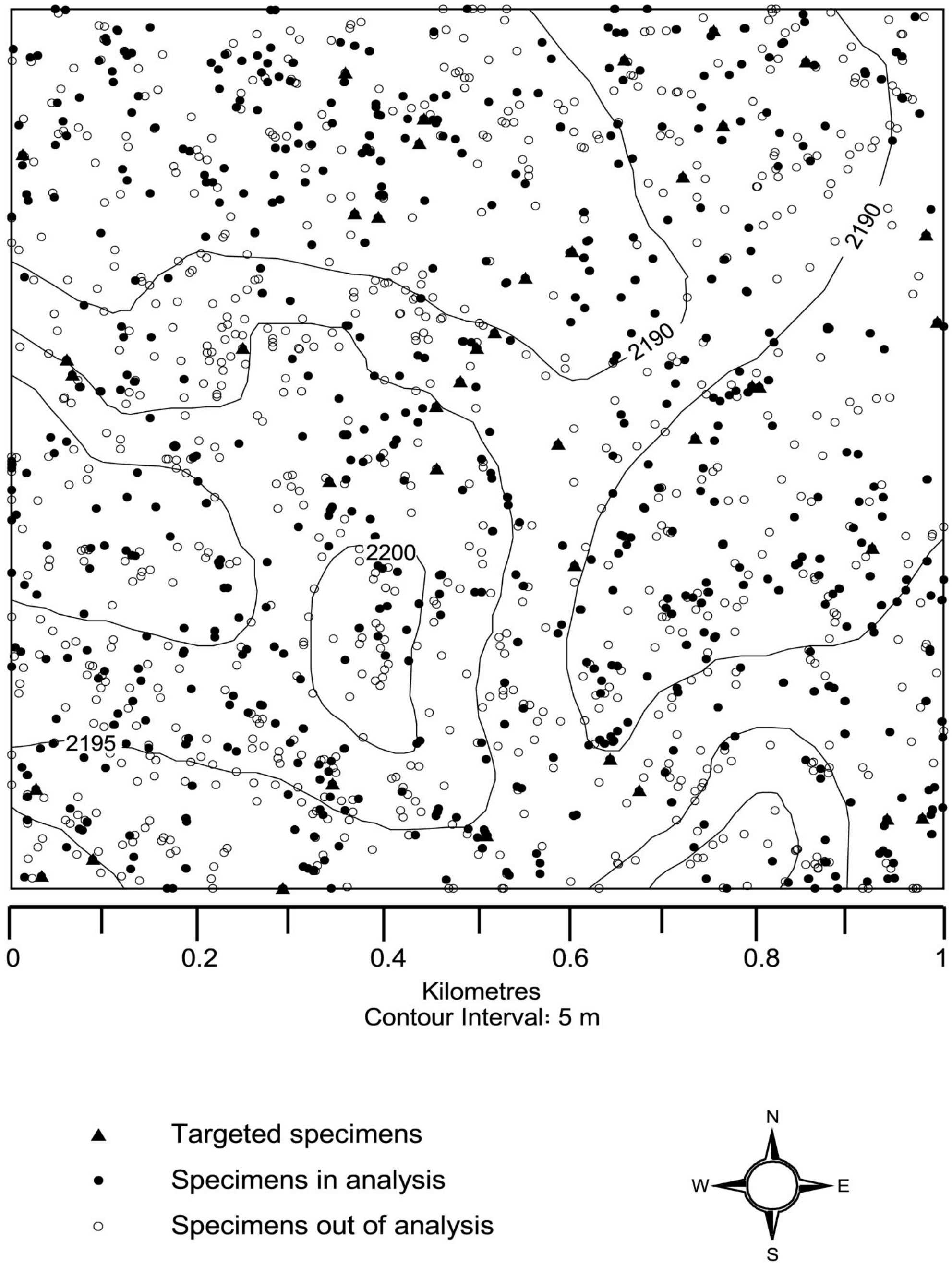
Results

Collection summary

A total of 1479 fire-scarred trees were documented and mapped (Fig. 3), and sections were collected from 1246 (84%) of these trees. Of the 233 (16%) trees from which specimens were not collected, 189 (13%) had a high level of decay preventing us from collecting a viable specimen. Specimens were not collected from the remaining 44 (3%) because they were young (<100 years) living trees with one scar and were clustered with other living recently single-scarred trees that we sampled. The following percentages were computed based on the 1246 collected specimens. We were able to cross-date 777 (62%) specimens and identify their fire dates, 67 (5%) of which were from live trees. Of the 459 specimens (37%) that were collected but not cross-dated, 303 (24%) contained an injury that was not necessarily a fire scar. The remaining 156 (13%) had clearly visible fire scars, but we failed to cross-date them because of decay and short or complacent ring series. Ten (0.8%) of the collected specimens went missing, either in the transfer from the field or during woodshop operations. Six hundred and forty-eight (52%) of the collected specimens had cross-dated fire scars within the period of analysis (1682–1881) and were used in the analysis described below.

During the period of analysis the percentage of recording trees varied between 16% and nearly 100% of the total sample depth (Fig. 4). The total number of trees sampled peaked between 1725 and 1750, whereas the sample depth (number of recording specimens) was greatest between 1800 and 1820. Both the total number of trees sampled and sample depth declined sharply near the end of the period of analysis. The fires that scarred more than 25% of the recording trees were clustered together in time. Larger fires and few small fires occurred between 1784 and 1813, with longer fire intervals

Fig. 3. Study area map of fire-scarred tree locations.



occurring between 1784–1788 and 1788–1794. Conversely, fires scarring fewer trees occurred almost annually from 1831 to 1850, but few large fires were recorded during this period. A notable lack of fires was recorded between 1873 and 1881.

Sampling method

Composite MFIs of all the sampling methods were within 2.18 years of each other and within 2.75 years of the census (Table 1).

Census and targeted sample

The census with no filter represented the maximum possible number of fire dates, so the census MFI was the shortest computed in this study. The targeted sample was slightly longer than the census with no filter and a 10% filter, but slightly shorter than the census with a 25% filter (Table 1). The targeted sample captured the specimen with the most scars in the whole study and included many more highly scarred specimens than the other samples. The standard deviation (SD) of the census with no filter was 0.7 years with fire intervals ranging from 1 to 4 years. The targeted sample SD was 1.1 years with fire intervals from 1 to 6 years.

Sample size (random samples) and area sampled

Mean MFI decreased towards an asymptote as the sample size increased for no filter and a 10% filter, while the mean MFI filtered at 25% remained fairly constant (Fig. 5*a*). The

Fig. 4. Composite fire history of all fire scarred trees (1682–1881) showing the number of trees scarred, sample depth (number of recording trees susceptible to fire scarring), and total number of fire-scarred trees present per year. The 10% and 25% filter lines indicate the minimum number of scarred trees required in that year to be included in the filter composite.

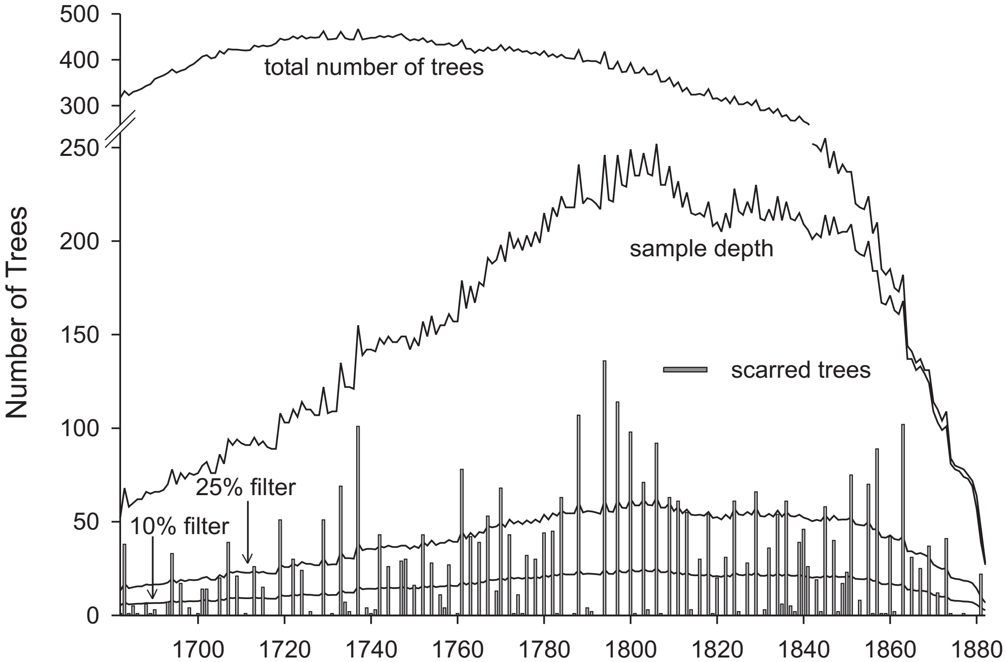


Table 1. Mean fire intervals (years) for all sampling methods at Centennial Forest, northern Arizona (1682–1881).

(A) Census and targeted sampling methods.

Filter	Census	Targeted
All	1.66	2.23
10%	2.83	3.00
25%	6.00	5.43

(B) Random sample size method.

	Sample size (no. of trees)							
Filter	10	20	30	40	50	60	70	80
All	4.41	3.71	3.05	2.82	2.43	2.58	2.46	2.28
10%	4.41	3.82	3.21	3.05	2.98	3.11	3.16	3.14
25%	4.94	5.87	6.00	6.25	6.84	5.95	6.28	5.66

(C) Sampled area method.

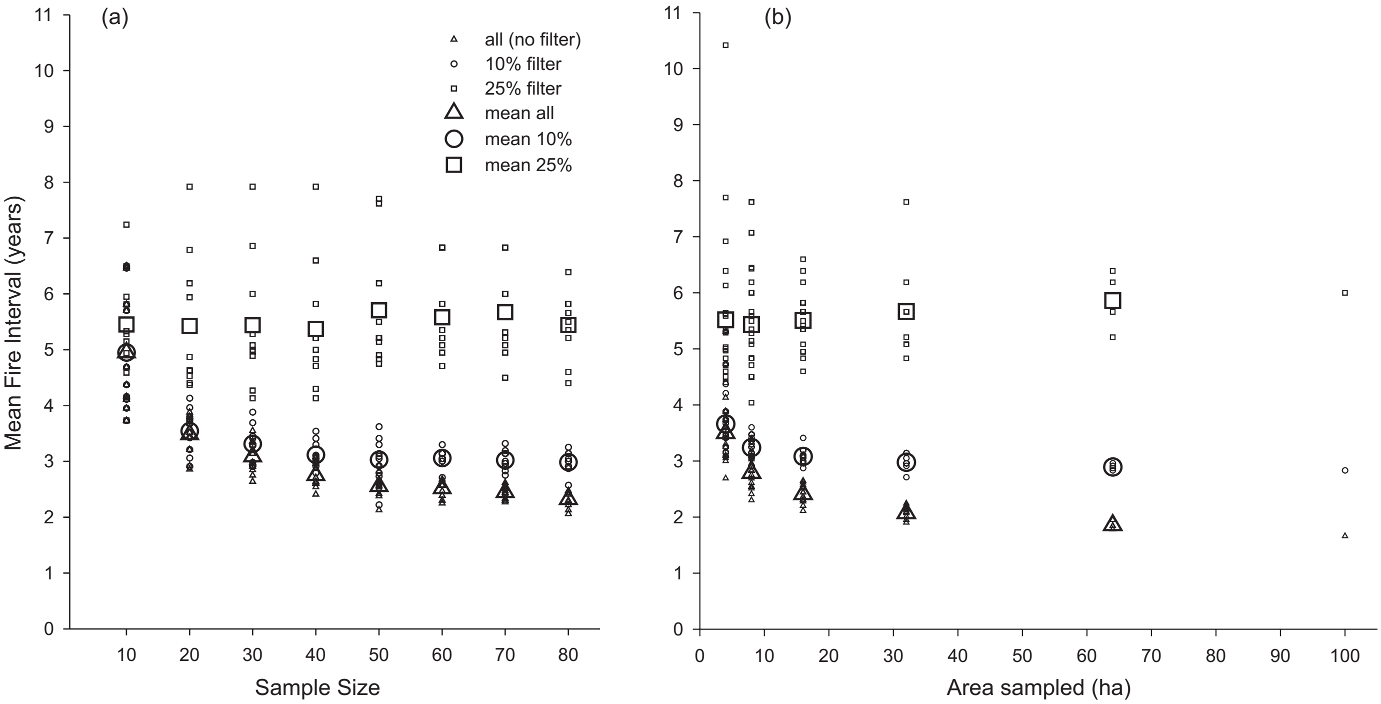
	Sampled area (no. of cells)				
Filter	1	2	4	9	16
All	3.49	2.79	2.41	2.07	1.85
10%	3.66	3.24	3.08	2.98	2.89
25%	5.52	5.44	5.51	5.67	5.86

(D) Grid-based method.*

	Search radius (m)				Checkerboard	
Filter	20	40†	40‡	60	Black	White
All	3.52	3.13	2.54	2.28	2.28	2.54
10%	3.52	3.4	2.91	2.71	2.87	2.96
25%	4.87	5.24	6	6.39	4.71	4.95
n	29	31	67	39	44	35
Specimens per plot	all (1–4)	1	2	1	1	1

Note: Each cell in the "sampled area" category was 200 m × 200 m.
*See Fig. 2 for grid layout maps.
†One specimen per plot.
‡Two specimens per plot.

Fig. 5. Mean fire intervals (MFIs) (years) from several samples showing the effect of (*a*) sample size and (*b*) area sampled. (*a*) Random samples of different sizes were used to test the effect of sample size. Ten samples were taken per unique combination of sample size and filter level (MFIs displayed by the small shapes), and the means of the sample MFIs are shown by the large shapes. (*b*) MFIs were computed for all combinations of each area category (4, 8, 16, 32, 64, and 100 ha) and are displayed by the small shapes.



variability within each set of 10 runs of each unique combination of sample size and filter level increased as the filter level increased. These trends were similar for area sampled (Fig. 5*b*). The 25% filtered means of MFI were very similar between small and large areas sampled. The MFIs decreased as area sampled increased for the less restrictive filters. Within the different samples of the same-sized area, variability was highest in the 25% filter MFIs and lower for the less restrictive filters.

As the sample size increased, the variability of fire intervals within a single sample decreased. SD of fire intervals in the random data sets of at least 40 trees with no filter ranged from 1.1 to 1.87 years. The maximum fire intervals in the same data sets ranged from 7 to 12 years. The minimum fire interval for these data sets was 1 year. Variability of SD and ranges of fire intervals increased with more restrictive filters.

Grid-based samples

In the samples based on grid 1, the longest MFIs resulted from the 20 and 40 m search radii, where the sample size was about 30 trees. The other two samples, one with a larger search radii and both with bigger sample sizes, resulted in shorter MFIs. The MFIs from the black or white cells in grid 2 were similar (Table 1).

Point and origin-to-first-scar intervals

The census composite MFIs with filters varied between 1.66 and 6 years. The variability of point MFIs was much greater, with a mean of 12 years and range from 2 to 133 years (Fig. 6). The targeted point MFI was 1.2 times shorter than the point MFI from the census data.

The average OS interval was 101.5 years. The frequency in the shorter intervals of the OS distribution increased until the 61–80 interval class, then declined (Fig. 7). Thirty-six percent of all specimens with pith scarred before age 70; 19.5% of specimens with pith scarred before age 50. The distance from a tree with an OS interval to its nearest neighbor scarred within the tree's OS interval ranged from 1 to 72 m, averaging 25 m. Of the fire-scarred neighbors, 87% were the closest recording neighbor to the tree in question. Many trees that scarred later in their lives remained unscarred during the most extensive fire years, including 1737 and 1794 (Fig. 8).

Discussion

Are targeting and other methods of sampling (random, grid based) accurate?

To answer this question, we needed to establish criteria to assess how well the sample MFIs represented the census MFI. We considered three different levels at which to determine similarity. The most restrictive criterion would be a 95% confidence interval (CI) of the census fire intervals. Because the range of fire intervals was only 1 to 4 years in the census, the 95% CI is very small, 0.13 years (= 47 days), so the threshold for similarity would be 1.79 years. This narrow threshold, met by none of the sampling methods, is not reasonable in terms of fire-scar formation and ecological or management considerations, as discussed below. A threshold at the other extreme would include MFIs less than 25 years, the maximum interval considered to represent a frequent fire regime (Pyne et al. 1996). This assessment would lead to the conclusion that all methods of sampling

Fig. 6. Distribution of individual fire intervals for the census data with three levels of filtering and the point (per tree) fire intervals. The 25th and 75th percentiles are denoted by either side of the shaded boxes with the median (vertical line) separating them. The whiskers extend to the 5th and 95th percentiles. The black dots are all the extreme values, and the white diamonds indicate the mean for each sample.

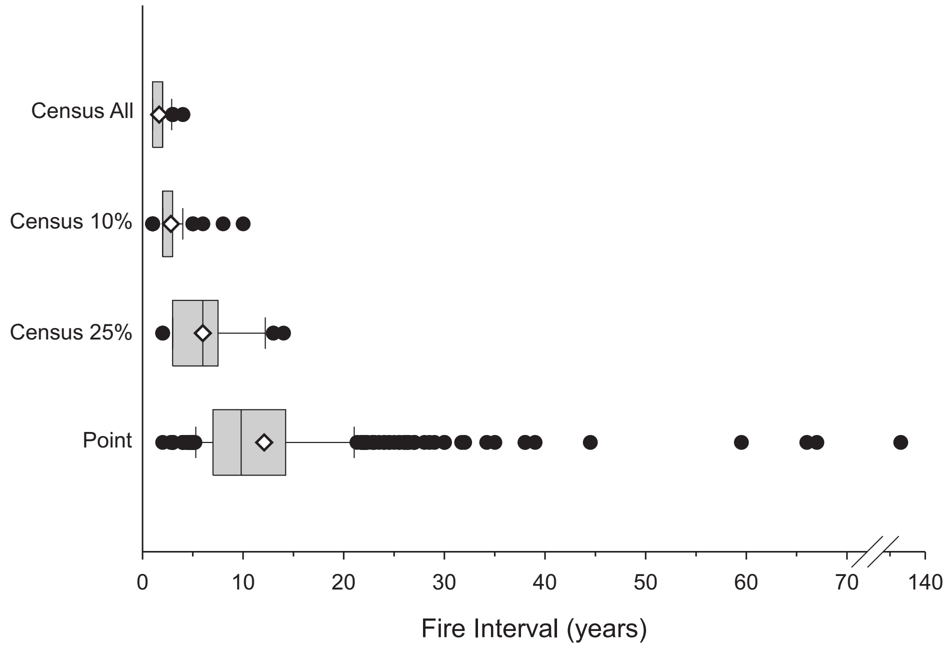
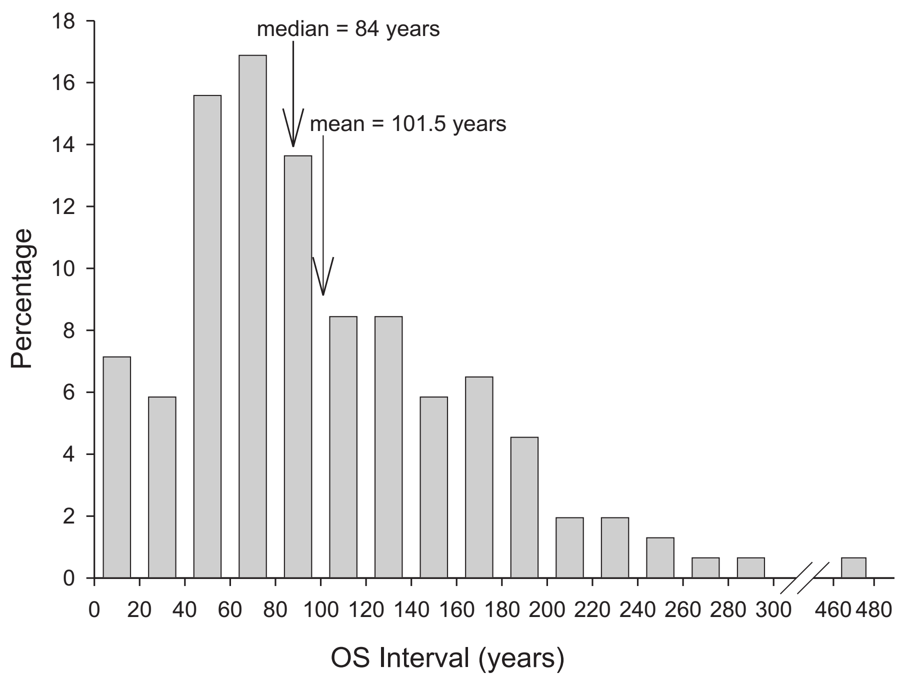


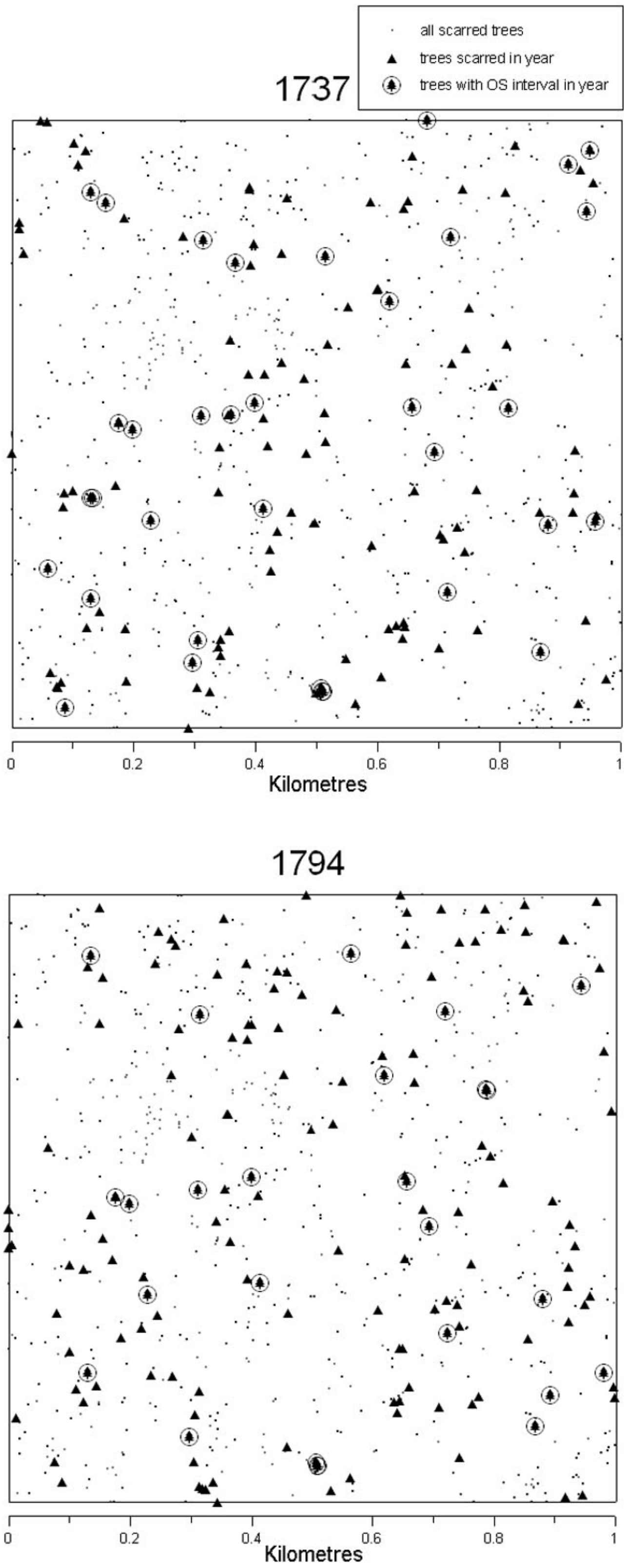
Fig. 7. Origin-to-first-scar (OS) interval for all trees with piths. This interval indicates the age of a tree when it received its first visible fire scar.



are adequate representations of the fire frequency, but the 25-year window is excessively broad for analysis of either ecological relationships or management practices. Between these extremes, we established the following reliability standard: all MFIs within 1 year of the census MFI are similar enough to represent the true fire frequency. One year is a reasonable threshold because it approximates the minimum temporal resolution (annual) at which fire-history data are recorded. If a different threshold is required for some analysis, the data in Table 1 can be compared against whatever criterion is deemed appropriate.

When the 1-year criterion is used, the threshold for samples is 2.66 years. Samples with MFIs that are similar to that of the census under the 1-year criterion include the targeted sample, random samples of at least 50 specimens, areas of at least 16 ha, and grid-based samples with at least 35 speci-

Fig. 8. Trees with origin-to-first-scar (OS) intervals occurred in close proximity to scarred trees during two major fire years, 1737 and 1794. OS trees are only mapped here if the OS interval overlaps the fire year shown.



mens. Accepting that a 1-year threshold is adequate, we conclude that all sampling methods tested in this study, given a sufficient sample size, will result in an accurate estimate of the true fire frequency.

How do sample size, area sampled, and filtering affect MFI estimates?

There is a threshold at which little new information is gained with additional specimens, as shown in this study (approximately 50 randomly sampled specimens) and elsewhere (Falk and Swetnam 2003; Stephens et al. 2003). However, smaller samples of targeted or grid-based specimens had similar results because we selected the specimens with the most fire dates in those samples. Swetnam and Baisan (1996) were correct that in a frequent surface fire regime, random sampling requires a larger sample size than targeted sampling to accumulate the same amount of fire-history data.

Sample size and area sampled are linked. The sample sizes in the 4 and 8 ha areas fell below the 50 specimen threshold, so the fact that MFIs in the 4 and 8 ha samples, 3.49 and 2.79 years, respectively, were longer than MFIs of larger areas may be a product of the small sample size, not necessarily a factor of area sampled. The same relationship was present in the 20 and 40 m samples, n = 29 and 31, respectively, based on grid 1. Even though there appeared to be an inverse relationship between search radius and MFI, the fact that MFIs in the 20 and 40 m samples, 3.52 and 3.13 years, respectively, were longer than MFIs for greater search radii may also be a function of sample size, not necessarily search radius.

The MFIs with a 25% filter were remarkably consistent as sample size or area sampled increased. This means that the large fires were captured with fewer specimens in a smaller area, but smaller fires continued to be discovered with more samples over a larger area. Therefore, if there were concern that fire-scar methods give undue weight towards small fires (Minnich et al. 2000; Baker and Ehle 2001), the use of the 25% filter should provide a stable basis for comparison even if the MFI calculated with all fires tends to vary with sample size and area.

Variability, for example, ranges and standard deviations of fire intervals, was not the focus of this study but it may have important ecological implications. The range (1 to 12 years) and SD (1.1 to 1.87 years) of fire intervals in random data sets of at least 40 specimens suggest that fire recurrence fluctuated from annual burning to fire-free gaps at decadal scales. Ecologically, this mosaic of fire frequencies allows for shifting patterns of understory vegetation, tree regeneration, and perhaps changes across a variety of trophic levels (e.g., Provencher et al. 2003). Large-scale climatic teleconnections have been suggested as a possible cause for a pattern of shifts in fire frequency and synchrony noted in Mexico (Stephens et al. 2003), Colorado, USA, and Patagonia, Argentina (Veblen and Kitzberger 2002), and elsewhere in North America (Swetnam and Betancourt 1990; Grissino-Mayer 1995; Grissino-Mayer and Swetnam 2000; Swetnam and Baisan 2003). We found a period of large synchronous fires between 1784 and 1813 and very frequent fires from 1831 to 1850, similar to the frequent fires that occurred from 1850 to 1865 at Chimney Spring (Dieterich 1980).

Should OS intervals, "bracketed" intervals, and point MFIs be included in fire history?

Two arguments were made in the literature that attempted to explain why most trees are older than 50 when they scar for the first time: (1) fires killed the young trees instead of scarring them, leaving no lasting evidence of the fire's presence, and (2) fire was absent around the seedling during its establishment (Baker 1989; Gutsell and Johnson 1996; Keeley and Stephenson 2000; Baker and Ehle 2001). While these are both logical explanations, a third possible conclusion missing from this list is that fire was present, but failed to leave a scar. The maps of widespread fire years (Fig. 8) show that it is highly likely that most trees experienced more than one fire without scarring, especially because some scarred trees are only 1 m away from a tree during its unscarred OS interval. The distribution of OS intervals for the Centennial Forest was similar to the distribution of OS intervals found for ponderosa pine growing in more northerly portions of the Rockies, except that measures of central tendency there were approximately 20 years longer (Baker and Ehle 2001).

The pith of a tree is not a surrogate for fire occurrence, especially in frequent fire regimes where postfire regeneration does not typically occur in even-aged cohorts (Mast et al. 1999). Baker (1989) noted that assuming that the OS interval is the same as a fire interval would be incorrect, though Baker and Ehle (2001, p. 1212) claimed that omitting the OS interval is an "error…since it estimates a real fire interval…" Furthermore, the OS interval is calculated to the first detectable fire scar; overestimation of the OS interval is likely considering the possibility of the first scar being burned away, broken, or decayed over time. In sum, while it is true that the OS period on any given tree could potentially represent a fire-free period, there is clear evidence of burning adjacent to yet-unscarred trees in our results. Coupled with the fact that there is no inherent link between a fire date and pith date, we suggest that the addition of OS intervals into the mathematical calculation of MFI, as proposed in the bracketing method of Baker and Ehle (2001), does not have a sound basis.

The point MFI, in contrast, is an interval between recorded fires (though scars can be burned away or decay, like all fire scars). They may be a useful metric of the maximum fire interval at the scale of individual trees, which recorded fires, on average, every 12 years in this study. This is longer than all the composite MFIs regardless of sampling method, yet still indicative of a frequent fire regime (<25 years). Unlike a composite of fire years, the point MFI is sensitive to the quality (number of fires) per specimen. Targeted sampling yields higher quality samples that may overestimate the point MFI and may not be appropriate for quantifying the maximum fire interval. In this study, the targeted point MFI was overestimated by a factor of 1.2 (<2 years). Although this was a small difference, if the point MFI is used to represent the maximum fire interval in other studies, perhaps a random sample of fire-scarred remnants should be used to quantify this interval.

The bracketing methods proposed by Baker and Ehle (2001) are an amalgam of useful quantitative techniques, such as filtering and point fire intervals, mixed together with the OS interval (not a fire interval) and arbitrary assertions about accuracy. For example, Baker and Ehle (2001) argued that "targeting likely decreases the mean composite FI by a factor of two to three times", a guess that was not supported in our data (target MFI = 2.33 years, census MFI = 1.66 years). Quantitative values and arbitrary estimates were lumped together by Baker and Ehle (2001) to arrive at bracketed MFI estimates of 22–308 years. Instead of following this approach, we suggest that fire-scar data be interpreted only with quantitative fire-interval data, including the unfiltered and filtered composite and point MFI values, together with descriptive statistics of variability.

How does this study relate to other fire-history studies in the region and the management of these forests?

The previously published fire-history studies from the region of our study site also reported high fire frequency before Euro-American settlement. Dieterich (1980) reported MFIs of 2.4 years at Chimney Spring and 1.8 years at Limestone Flat; sites on the San Francisco Peaks had an MFI of 5.2 years (Heinlein 1996); Fulé et al. (1997) reported a 3.7 year MFI for Camp Navajo (Fig. 1). These studies also reported similar dates for the cessation of the frequent fire regime, from 1876 at Chimney Springs (Dieterich 1980) to 1883 at Camp Navajo (Fulé et al. 1997). All the northern Arizona studies fall within the range of results reported across the southwestern United States by Swetnam and Baisan (1996). Modern calibration studies (Farris et al. 2003; Fulé et al. 2003, Stephens et al. 2003) indicate that fire-scar data are consistent with independent fire records in reconstructing fire occurrence. Other lines of evidence from evolutionary ecology (Moore et al. 1999) and historical documentation (Cooper 1960) support the interpretations of ponderosa pine fire regimes in the southwestern United States obtained from numerous studies based on fire-scar analyses.

Our data are limited to a single case study of 1 km^2, though it is worth noting that the 648 cross-dated specimens are equivalent to more than half of the total of 1215 specimens at all 63 southwestern study sites reported by Swetnam and Baisan (1996). Because fire-scarred trees provide point estimates of fire occurrence, our sampling cannot resolve the uncertainty of burning between these points nor the interpretation of fire occurrence around never-scarred or yet-unscarred (OS) trees. However, we were able to completely resolve the concerns that targeting and other sampling methods may have an unknown effect on estimates of fire frequency at this study site. All methods of sampling compared in this study are accurate given a sufficient sample size: any 50 specimens would yield a fair estimate of the fire frequency. Targeting requires the smallest sample size, yields the same results as other sampling methods, and is likely to result in longer reliable records of fire.

In recent years, wildfires in the southwestern United Sates have dramatically increased in size and severity, resulting in undesirable ecological effects (Agee 1993; Kolb et al. 1994; Swetnam et al. 1999) and increased costs of suppression and rehabilitation (National Fire Plan 2004; GAO 2004). Recent legislation has encouraged thinning and prescribed burning (Healthy Forest Restoration Act 2003). Because management recommendations are based partially on historic forest conditions and fire frequencies, it is important to have information collected in such a way that accurately represents the true historic conditions, though even given perfect knowledge of the historic fire regime, it is unlikely that managers would implement equally frequent, widespread fires because of other constraints.

As long as managers recognize that fire-scar-based fire histories represent patterns of fires that burned in variable spatial and temporal patterns, with a mosaic of burn severities, including unburned areas within the fire perimeter, there is no evidence from our study site to support Baker and Ehle's (2003, p. 329) contention that "traditional measures [of fire history] are misleading or in error as sources of [restoration] information". In fact, targeting and every other sampling approach with ≥50 specimens were essentially indistinguishable from census data and even low sample sizes or small areas sampled were quite stable when data were filtered to the 25% scarred level. The high consistency achieved among sampling methods and the consistency between our data and southwestern ponderosa pine fire histories in general are strong evidence against the contention that repeated management burning "lacks a sound basis in science" (Baker and Ehle 2003, p. 330), at least in southwestern ponderosa forests and perhaps more broadly in the range of ponderosa and related pines.

Acknowledgements

Thanks to the Centennial Forest of Northern Arizona University, J.J. Smith, and M. Wagner. N. Yepa, E. Hume, D. Fairbairn, and D. Normandin were instrumental in collecting and analyzing samples. J. Crouse helped prepare the maps. T. Swetnam, C. Hull Sieg, R. King, T. Kolb, H. Grissino-Mayer, P. Brown, and an anonymous reviewer provided comments to improve the manuscript. Research funding came from the Mission Research program of the School of Forestry and the Ecological Restoration Institute, Northern Arizona University.

References

Abella, S.R. 2005. Vegetational and environmental gradients on an Arizona ponderosa pine landscape: implications for ecological restoration. Ph.D. dissertation, Northern Arizona University, Flagstaff, Ariz.

Agee, J.K. 1993. Fire ecology of Pacific Northwest forests. Island Press, Washington, D.C.

Arno, S.F., and Sneck, K.M. 1977. A method for determining fire history in coniferous forests of the Mountain West. USDA For. Serv. Gen. Tech. Rep. INT-42.

Arno, S.F., Scott, J.H., and Hartwell, M.G. 1995. Age-class structure of old-growth ponderosa pine/Douglas-fir stands and its relationship to fire history. USDA For. Serv. Res. Pap. INT-481.

Baker, W.L. 1989. Effect of scale and spatial heterogeneity on fire-interval distributions. Can. J. For. Res. **19**: 700–706.

Baker, W.L., and Ehle, D.S. 2001. Uncertainty in surface-fire history: the case of ponderosa pine forests in the western United States. Can. J. For. Res. **31**: 1205–1226.

Baker, W.L., and Ehle, D.S. 2003. Uncertainty in fire history and restoration of ponderosa pine forests in the western United States. *In* Fire, Fuel Treatments, and Ecological Restoration: Conference Proceedings, 16–18 April 2002, Fort Collins, Colo. *Edited by* P.N. Omi and L.A. Joyce. USDA For. Serv. Pap. RMRS-29. pp. 319–333.

Cooper, C.F. 1960. Changes in vegetation, structure and growth of southwestern pine forests since white settlement. Ecol. Monogr. **30**: 129–164.

Davis, K. 1987. Fire history of Walnut Canyon National Monument. Unpublished report to the National Park Service, Flagstaff Area National Monument files.

Dieterich, J.H. 1980. The composite fire interval — a tool for more accurate interpretations of fire history. *In* Proceedings of the Fire History Workshop, 20–24 October 1980, Tucson, Ariz. *Edited by* M.A. Stokes and J.H. Dieterich. USDA For. Serv. Gen. Tech. Rep. RM-81. pp. 8–14.

Dieterich, J.H., and Swetnam, T.W. 1984. Dendrochronology of a fire-scarred ponderosa pine. For. Sci. **30**: 238–247.

Falk, D.A., and Swetnam, T.W. 2003. Scaling rules and probability models for surface fire regimes in ponderosa pine forests. *In* Fire, Fuel Treatments, and Ecological Restoration: Conference Proceedings. 16–18 April 2002, Fort Collins, Colo. *Edited by* P.N. Omi and L.A. Joyce. USDA For. Serv. Pap. RMRS-29. pp. 301–317.

Fall, J.G. 1998. Reconstructing the historical frequency of fire: a modeling approach to developing and testing methods. M.S. thesis, Simon Fraser University, Burnaby, B.C.

Farris, C.A., Baisan, C.H., and Swetnam, T.W. 2003. A comparison of fire regime characteristics reconstructed from fire scar data and mapped fires in a frequently burned Arizona wilderness. Presented at the Second International Wildland Fire Ecology and Fire Management Congress, 16–20 November 2003, Orlando, Fla. Abstract available from http://ams.confex.com/ams/FIRE2003/techprogram/paper_65930.htm.

Fulé, P.Z., Covington, W.W., and Moore, M.M. 1997. Determining reference conditions for ecosystem management of southwestern ponderosa pine forests. Ecol. Appl. **7**: 895–908.

Fulé, P.Z., Heinlein, T.A., Covington, W.W., and Moore, M.M. 2003. Assessing fire regimes on Grand Canyon landscapes with fire-scar and fire-record data. Int. J. Wildland Fire, **12**: 129–145.

GAO. 2004. Wildland fires: Forest Service and BLM need better information and a systematic approach for assessing the risks of environmental effects. Government Accountability Office, Washington, D.C. Rep. GAO-04-705.

Graybill, D.A. 1987. Gus Pearson, Arizona ring width chronology. Tree-ring search engine. Available from http://www.ngdc.noaa.gov/paleo/treering.html [accessed 1 July 2003].

Grissino-Mayer, H.D. 1995. Tree-ring reconstructions of climate and fire history at El Malpais National Monument, New Mexico. Ph.D. dissertation, University of Arizona, Tucson, Ariz.

Grissino-Mayer, H.D. 1999. Modeling fire interval data from the American Southwest with the Weibull distribution. Int. J. Wildland Fire, **9**: 37–50.

Grissino-Mayer, H.D., and Swetnam, T.W. 2000. Century-scale climate forcing of fire regimes in the American Southwest. Holocene, **10**: 207–214.

Grissino-Mayer, H.D., Romme, W.H., Floyd, M.L., and Hanna, D.D. 2004. Climatic and human influences on fire regimes of the southern San Juan Mountains, Colorado, USA. Ecology, **85**: 1708–1724.

Gutsell, S.L., and Johnson, E.A. 1996. How fire scars are formed: coupling a disturbance process to its ecological effect. Can. J. For. Res. **26**: 166–174.

Healthy Forest Restoration Act. 2003. H.R. 1904 [online]. Available from http://www.theorator.com/bills108/h1904.html [cited 6 December 2004].

Heinlein, T.A. 1996. Fire regimes and forest structure in lower mixed conifer forests: San Francisco Peaks, Arizona. M.S. thesis, Northern Arizona University, Flagstaff, Ariz.

Heyerdahl, E.K. 1997. Spatial and temporal variation in historical fire regimes of the Blue Mountains, Oregon and Washington: the influence of climate. Ph.D. dissertation, University of Washington, Seattle, Wash.

Heyerdahl, E.K., and McKay, S.J. 2001. Condition of live fire-scarred ponderosa pine trees six years after removing partial cross sections. Tree-Ring Res. **57**: 131–139.

Holmes, R.L. 1983. Computer-assisted quality control in tree-ring dating and measurement. Tree-Ring Bull. **43**: 69–78.

Johnson, E.A., and Gutsell, S.L. 1994. Fire frequency models, methods and interpretations. Adv. Ecol. Res. **25**: 239–287.

Keeley, J.E., and Stephenson, N.L. 2000. Restoring natural fire regimes to the Sierra Nevada in an era of global change. *In* Wilderness Science in a Time of Change Conference. Vol. 5. Wilderness Ecosystems, Threats, and Management, 23–27 May 1999, Missoula, Mont. *Edited by* D.N. Cole, S.F. McCool, W.T. Borrie, and J.O. O'Loughlin. USDA For. Serv. Pap. RMRS-15. pp. 255–265.

Kolb, T.E., Wagner, M.R., and Covington, W.W. 1994. Concepts of forest health: utilitarian and ecosystem perspectives. J. For. **92**: 10–15.

Mast, J.N., Fulé, P.Z., Moore, M.M., Covington, W.W., and Waltz, A. 1999. Restoration of presettlement age structure of an Arizona ponderosa pine forest. Ecol. Appl. **9**: 228–239.

McBride, J.R. 1983. Analysis of tree rings and fire scars to establish fire history. Tree-Ring Bull. **42**: 51–67.

Miller, G., Ambrose, N., Boness, P., Reher, D., Robertson, G., Scalzone, K., Steinke, R., and Subirge, T. 1995. Terrestrial ecosystem survey. USDA Forest Service, Southwestern Region 3, Albuquerque, N.M.

Minnich, R.A., Barbour, M.G., Burk, J.H., and Sosa-Ramirez, J. 2000. Californian mixed-conifer forests under unmanaged fire regimes in the Sierra San Pedro Martir, Baja California, Mexico. J. Biogeogr. **27**: 105–129.

Moore, M.M., Covington, W.W., and Fulé, P.Z. 1999. Reference conditions and ecological restoration: a southwestern ponderosa pine perspective. Ecol. Appl. **9**: 1266–1277.

National Fire Plan. 2004. Large fire suppression costs, strategies for cost management [online]. Available from http://www.fireplan.gov/resources/annual_report.html [cited 6 December 2004].

Provencher, L., Litt, A.R., and Gordon, D.R. 2003. Predictors of species richness in northwest Florida longleaf pine sandhills. Conserv. Biol. **17**: 1660–1671.

Pyne, S.J., Andrews, P.L., and Laven, R.D. 1996. Introduction to wildland fire, 2nd ed. John Wiley and Sons, Inc., New York.

Stephens, S.L., Skinner, C.N., and Gill, S.J. 2003. Dendrochronology-based fire history of Jeffrey pine – mixed conifer forests in the Sierra San Pedro Martir, Mexico. Can. J. For. Res. **33**: 1090–1101.

Stokes, M.A., and Smiley, T.L. 1968. An introduction to tree-ring dating. University of Chicago, Chicago, Ill.

Swetnam, T.W., and Baisan, C.H. 1996. Historical fire regime patterns in the southwestern United States since AD 1700. *In* Fire Effects in Southwestern Forests: Proceedings of the Second La Mesa Fire Symposium, 29–31 March 1994, Los Alamos, N.M. *Edited by* C.D. Allen. USDA For. Serv. Gen. Tech. Rep. RM-286. pp. 11–32.

Swetnam, T.W., and Baisan, C.H. 2003. Tree-ring reconstructions of fire and climate history in the Sierra Nevada and southwestern United States. *In* Fire and climate change in temperate ecosystems of the western Americas. *Edited by* T.T. Veblen, W.L Baker, G. Montenegro, and T.W. Swetnam. Springer, New York. pp. 158–195.

Swetnam, T.W., and Betancourt, J.L. 1990. Fire–Southern Oscillation relations in the southwestern United States. Science (Washington, D.C.), **249**: 1017–1020.

Swetnam, T.W., Wright, W.E., Caprio, A.C., Brown, P.M., and Baisan, C.H. 1990. Fire scar dates from Walnut Canyon National Monument, Arizona. Unpublished report to the National Park Service, Flagstaff Area National Monument files.

Swetnam, T.W., Allen, C.D., and Betancourt, J.L. 1999. Applied historical ecology: using the past to manage for the future. Ecol. Appl. **9**: 1189–1206.

Veblen, T.T., and Kitzberger, T. 2002. Inter-hemispheric comparison of fire history: the Colorado front range, U.S.A., and the northern Patagonian Andes, Argentina. Plant Ecol. **163**: 187–207.

Western Regional Climate Center. 2004. Western U.S. historical climate summaries [online]. Available from http://www.wrcc.dri.edu/cgi-bin/cliMAIN.pl?azflag [cited 13 April 2005].