

# Modeling Forest Canopy Structure and Density by Combining Point Quadrat Sampling and Survival Analysis

Daniel S. Maynard, Mark J. Ducey, Russell G. Congalton, and Joel Hartter

**Abstract:** Point quadrat sampling has been used relatively infrequently for modeling canopy structure and density, primarily because of the large number of sample points needed to obtain accurate estimates. We address these limitations by showing how point quadrat data are a form of time-to-event data, analogous to what are commonly observed in biomedical studies. This equivalence allows for point quadrat data to be analyzed using existing survival analysis methods. We illustrate the usefulness of this relationship by analyzing data from a field study conducted in northeast Oregon. Within each of 60 forest plots, we obtained canopy-height measurements using a handheld laser rangefinder, and we used a survival-based regression model to estimate canopy profiles and leaf area indices via the Weibull hazard function. The resulting survival-based estimates of canopy density and structure appeared robust to sample size limitations, whereas the relatively small number of samples per plot led to an apparent underestimation of canopy density by the traditional point quadrat estimator. Overall, the incorporation of survival analysis methods and point quadrat sampling greatly increases the usefulness of this sampling method, resulting in an efficient tool for quickly assessing the structure of forest canopies. *FOR. SCI.* 59(6):681–692.

**Keywords:** leaf area index, vertical structure, forest sampling, regression, variance estimation

THE structure and density of forest canopies are important indicators of forest productivity and function. Canopy characteristics are strong indicators of individual tree growth, vigor, and overall stand productivity (Waring et al. 1981, Mitchell et al. 1983, Waring 1983, Maguire and Kanaskie 2002). Canopies are responsible for energy production, carbon sequestration, and water transfer (Ford and Deans 1978, Hollinger 1989, Baldocchi et al. 2002, Baldocchi and Harley 2006). They regulate the quantity of light and water that reach the forest floor, thereby influencing temperature, humidity, and moisture (Seidel et al. 2011). These effects can alter understory vegetation patterns and corresponding wildlife populations (Hayes et al. 1997, DeMaynadier and Hunter 1999, Jennings et al. 1999, Martens et al. 2000, Wilson et al. 2007). Accurate estimation of canopy structure is integral for making responsible and effective stand-level management decisions regarding the health, productivity, habitat suitability, and ecosystem processes of forests.

Numerous ground-based methods exist for assessing canopy structure and canopy density, although most are subject to one or more limitations: they can be expensive, time-consuming, and difficult to implement, or they can yield data that are imprecise or hard to interpret (Radtke and Bolstad 2001, Jonckheere et al. 2004, Côté et al. 2011, Seidel et al. 2011, 2012, Zhao et al. 2011). Common techniques for estimating canopy structure include stratified clipping and scaffolding (Fujimori 1971, Fukushima et al.

1998), litter traps (Ovington 1963, Aber 1979), inclined point quadrat methods, and optical or photographic methods (Levy and Madden 1933, Wilson 1960, MacArthur and Horn 1969, Aber 1979, Rich 1990). More recently, ground-based optical point quadrat methods such as light detection and ranging (LiDAR) and laser-point sampling have been used with increasing frequency (Vanderbilt et al. 1979, Radtke and Bolstad 2001, Lovell et al. 2003, Parker et al. 2004, Coops et al. 2007).

In this article, we build on the point quadrat and laser point quadrat sampling methods of Levy and Madden (1933), MacArthur and Horn (1969), and Radtke and Bolstad (2001) by illustrating how laser point quadrat sampling is analogous to the field of survival analysis as used in biomedical and technology research. In doing so, we hope to alleviate many of the limitations that are common to ground-based canopy sampling methods. The objectives of this article are to demonstrate the equivalence between survival analysis and point quadrat analysis, to provide an example that illustrates the potential usefulness of this relationship, and to investigate how survival-based regression analysis can be used to overcome limitations present in traditional point quadrat sampling.

First, we give a brief review of existing point quadrat sampling methods, followed by an overview of survival analysis methods. We then outline the previously undescribed relationship between survival analysis and point quadrat sampling. To investigate the feasibility and efficacy

Manuscript received July 26, 2012; accepted February 8, 2013; published online April 4, 2013; <http://dx.doi.org/10.5849/forsci.12-086>.

Daniel S. Maynard ([daniel.maynard@yale.edu](mailto:daniel.maynard@yale.edu)), Yale School of Forestry & Environmental Studies, Yale University, New Haven, CT. Mark J. Ducey ([mjducey@cisunix.unh.edu](mailto:mjducey@cisunix.unh.edu)), Department of Natural Resources and the Environment, University of New Hampshire. Russell G. Congalton ([russ.congalton@unh.edu](mailto:russ.congalton@unh.edu)), Department of Natural Resources and the Environment, University of New Hampshire. Joel Hartter ([joel.hartter@unh.edu](mailto:joel.hartter@unh.edu)), Department of Natural Resources and the Environment, University of New Hampshire.

Acknowledgments: Funding for this research was provided in part by the Disaster Resilience for Rural Communities Program through the USDA National Institute of Food and Agriculture (Award 2010-67023-21705) and by the New Hampshire Agricultural Experiment Station.

Copyright © 2013 by the Society of American Foresters.

of applying regression-based survival analysis techniques to point laser quadrat data, we conducted a field study in northeast Oregon. The results and conclusions from this study are reported, along with limitations and suggestions for further research on the use of survival analysis for point quadrat data.

## Point Quadrat Sampling

The primary variables of interest in point quadrat sampling are leaf area index, canopy density, and canopy profile. Leaf area index (LAI) can be defined numerous ways: as the horizontally projected leaf area (Ross 1981, Bolstad and Gower 1990); as the total one-sided leaf area per unit ground area (Watson 1947); or as one-half of the total leaf area per unit ground area (Chen and Black 1992a). Following the tradition of Levy and Madden (1933), MacArthur and Horn (1969), and Radtke and Bolstad (2001), we do not account for the angular distribution of foliage in our analysis. Our use of the term “leaf area index” in this article therefore refers to horizontally projected leaf area per unit ground area. In addition, laser point quadrat sampling measures plant area index (projected plant area per unit ground area), rather than leaf area index, because the laser does not differentiate between foliage and woody material. Despite this limitation, we use the term “leaf area index” (LAI) rather than “plant area index” to coincide with traditional point quadrat sampling terminology. Possible ways of overcoming this limitation using survival analysis methods are addressed in the Discussion.

Chen and Black (1992a) point out that the horizontal projection for nonflat leaves can be problematic and has limited biological significance. Furthermore, the clumping of foliage due to the nonrandomness of shoot positions and the interdependence of leaf positions in the canopy can lead to underestimates of LAI (Chen and Black 1992b, Chen and Cihlar 1995, Radtke and Bolstad 2001). Although these limitations can affect the interpretability of the results, the purpose of the proposed methods is to further the applicability of point quadrat sampling by highlighting the similarities between survival analysis and existing point quadrat methods. Techniques for addressing these limitations have been discussed elsewhere (see Wilson 1960, 1963, 1965, Chen and Black 1992a, Chen and Cihlar 1995, Welles and Cohen 1996, Denison 1997, Groeneveld 1997, Radtke and Bolstad 2001).

### MacArthur and Horn Estimator

The original methods for conducting point quadrat sampling involved passing a needle (i.e., a sewing needle) vertically through the canopy and recording the number of contacts between the needle and foliage (Levy and Madden 1933). MacArthur and Horn (1969) expanded on this method by using a camera to measure height to first leaf. According to this scheme, a random point was selected, a tripod was centered over that location, and a camera was fixed to the tripod and sighted vertically up into the forest canopy. Using the focusing scale on the telephoto lens, distance to the first leaf at that point was measured. By

repeating this procedure across a random series of sample points they obtained a set of measurements of “height to first leaf.”

Using these measurements, MacArthur and Horn (1969) derived a method for estimating horizontally projected LAI. Let  $D(h)$  represent the foliage density, in terms of projected leaf area per unit volume, of an infinitesimally small segment of the canopy, centered at height  $h$  (this function is also known as the “leaf area density” function; Weiss 2004). Let  $\varphi(h)$  be the probability that there are no leaves over the first  $h$  meters. MacArthur and Horn (1969) derived the following relationship between  $D(h)$  and  $\varphi(h)$

$$D(h) = -\frac{d[\log\varphi(h)]}{dh}. \quad (1)$$

The LAI between any two heights  $h_1$  and  $h_2$  ( $h_1 < h_2$ ) is found by integrating the foliage density function between these endpoints. Using the above equality, we have that

$$\text{LAI}(h_1, h_2) = \int_{h_1}^{h_2} D(h)dh = \log\frac{\varphi(h_1)}{\varphi(h_2)}. \quad (2)$$

As  $\varphi(h)$  is not known, MacArthur and Horn (1969) suggest estimating it via the proportion of measurements that exceed  $h$ . Under this estimation procedure, Equation 2 can be estimated by

$$\text{LAI}(h_1, h_2) = \log\frac{n_{h_1}}{n_{h_2}}, \quad (3)$$

where  $\text{LAI}(h_1, h_2)$  is the estimated LAI between heights  $h_1$  and  $h_2$  and  $n_{h_1}$  and  $n_{h_2}$  are the number of measurements that exceeded these two heights, respectively. The MacArthur-Horn estimate of the LAI of the site is therefore equal to

$$\text{LAI}_{\text{MH}} = \log\frac{N}{n_0} = \log(g^{-1}), \quad (4)$$

where  $N$  is the total number of points,  $n_0$  is the number of points that did not intercept any foliage, and  $g$  is the gap fraction (i.e., the proportion of sample points that did not intercept any foliage).

### Laser Point Sampling—Radtke and Bolstad Estimator

Radtke and Bolstad (2001) modified the sampling procedure of MacArthur and Horn (1969) by using a laser rangefinder instead of a camera to measure the distance to first leaf. The laser rangefinder was attached to a tripod and a measurement was taken every 10–13 cm in a grid pattern throughout a  $13 \times 13$ -m plot within each stand. To analyze these data, the authors provided an alternative estimator to that proposed by MacArthur and Horn (1969). They divide the canopy into a hypothetical series of nonoverlapping intervals of variable width, such that only a single laser shot is intercepted within each interval. Under this assumption, the Radtke-Bolstad estimate of leaf area in interval  $i$  is equal to

$$q_i = \frac{1}{n_i}, \quad (5)$$

where  $n_i$  is the number of measurements that were intercepted in or above interval  $i$ .

To estimate the leaf area between two heights,  $h_L$  and  $h_U$  ( $h_L < h_U$ ), the leaf area estimates for the each interval between these two heights are summed

$$\text{LAI}(h_L, h_U) = \sum_{j: h_L \leq h_j < h_U} \frac{1}{n_j}. \quad (6)$$

Radtke and Bolstad (2001) show that this estimator is asymptotically equivalent to Equation 3 under a few simple constraints. To obtain canopy profiles, Radtke and Bolstad (2001) divided the canopy into 1-m height intervals and used Equation 6 to estimate and plot the foliage density within each interval.

## Survival Analysis

In biostatistics, whenever a subject is followed until a predefined event occurs, the resulting data are termed “failure time data” or “time-to-event data” (Kalbfleisch and Prentice 2002). The group of techniques and methods used to analyze failure time data is known as “survival analysis.” Failure time data occur most frequently in biomedical studies in which patients or persons of interest are followed for a period of time until a specific event occurs (e.g., death, adverse reaction, or clinical diagnosis). It is also commonly used in industrial applications, in which failure rates and life expectancies of specific components are of interest.

Survival analysis generally focuses on making inferences about two functions: the survivor function and the hazard function. The probability that the predefined event occurs after time  $t$  (i.e., the probability that the subject or component “survives” up to and including time  $t$ ) is given by the survivor function

$$S(t) = 1 - F(t) = P(T > t), \quad 0 < t < \infty. \quad (7)$$

where  $F(t)$  is the cumulative distribution function. This equation is sometimes also referred to as the “reliability function” (Kalbfleisch and Prentice 2002, Tableman et al. 2004). This is a monotonic decreasing function that equals 1 at time zero and approaches 0 as  $t$  approaches infinity.

In survival analysis, the hazard function specifies the instantaneous rate of failure at time  $t$ . It can be defined in terms of the survivor function

$$\begin{aligned} h(t) &= \lim_{\delta \rightarrow 0^+} \frac{P(t \leq T < t + \delta | T \geq t)}{\delta} \\ &= - \frac{d \log S(t)}{dt}. \end{aligned} \quad (8)$$

By taking the derivative of  $\log S(t)$ , the hazard function can likewise be defined in terms of the probability density function,  $f(t)$ , and the survivor function

$$h(t) = \frac{f(t)}{S(t)}. \quad (9)$$

The hazard function is often referred to as a hazard rate function, as it can be conceptualized as representing the failure rate at any given time. It follows that the cumulative hazard between any two time points  $t_1$  and  $t_2$  is given by

$$\int_{t_1}^{t_2} h(t) dt = \log \frac{S(t_1)}{S(t_2)}. \quad (10)$$

## Equivalence of Point Quadrat Sampling and Survival Analysis

Although the most common estimators for laser point quadrat analysis and its optical analogs have been known since the work of MacArthur and Horn (1969), the connection to survival analysis does not seem to have been recognized in the literature. There are a considerable number of statistical techniques associated with survival analysis that are useful in the analysis of laser or optical point quadrat data, including straightforward variance estimation, hypothesis testing, and the incorporation of covariates in the analysis.

To understand this equivalence between point quadrat sampling and survival analysis, we can imagine that each laser shot is “enrolled” at height zero and followed upward in height until the predefined event (i.e., leaf interception) occurs. Just as in the traditional survival setting, this results in time-to-event data. The difference is simply the reference scale: in traditional survival analysis, the subjects are followed across time until the first event occurs; in point quadrat sampling, the subjects (laser shots) are followed across distance until the first leaf intercept occurs. As evidenced by the similarities between Equations 1 and 8 and between Equations 2 and 10, this change from time to distance does not affect the underlying mathematics.

In point quadrat sampling,  $\varphi(h)$  is defined as the probability that there are no leaves over the first  $h$  meters. This is equivalent to the probability that a laser beam “survives” to height  $h$  without intercepting a leaf (or woody material), and it is identical to the survivor function. The MacArthur and Horn density function  $D(h)$  is also equivalent to the hazard function  $h(t)$ , as both are describing an instantaneous risk of failure, the former in terms of leaves per meter and the latter in terms of events per time. Therefore, the following congruencies hold

$$D(h) \cong h(t) \quad (11)$$

$$\varphi(h) \cong S(t). \quad (12)$$

Furthermore, MacArthur and Horn (1969) suggested estimating  $\varphi(h)$  via the proportion of sample points that exceeded  $h$ . In survival analysis, this is known as the empirical survival function (Tableman et al. 2004). The overall estimate of leaf area as suggested by MacArthur and Horn (1969) in Equation 4 is equivalent to the Kaplan-Meier product limit estimate of the cumulative hazard (Kaplan and Meier 1958), provided that there are no censored observations (such observations are common in failure-time studies, where the subject is removed or withdraws from the study before the outcome can be



observed; methods of incorporating censoring are addressed in the Discussion).

Using the results established by Kaplan and Meier (1958), we can obtain the previously unrecognized variance estimate for the MacArthur-Horn LAI estimate. Let  $(h_1, \dots, h_n)$  be the ordered heights at which a leaf was intercepted,  $d_i$  denote the total number of leaf intercepts at  $h_i$ , and  $n_i$  denote the total number of points that were intercepted at or above  $h_i$ . It follows that the variance is (Tableman et al. 2004)

$$\widehat{\text{var}}(\text{LAI}_{\text{MH}}) = \widehat{\text{var}}\left(\log\left(\frac{1}{g}\right)\right) = \sum_{i=1}^n \frac{d_i}{n_i(n_i - d_i)}. \quad (13)$$

Similarly, the variance of the Radtke-Bolstad LAI estimator as shown in Equation 6 was previously unknown. This estimator is actually equivalent to the Nelson-Aalen estimate of the cumulative hazard function taken between two fixed heights (Tableman et al. 2004), provided there are no two identical laser intercept heights (in practice, measurement error generally assures that no two heights measurements will be identical). The variance of Equation 6 is therefore equal to (Nelson 1972, Aalen 1978)

$$\widehat{\text{var}}(\text{LAI}(h_1, h_2)) = \sum_{j: h_1 \leq h_j < h_2} \frac{1}{n_j^2}. \quad (14)$$

Not only does survival analysis allow for variance estimation of traditional, nonparametric point quadrat estimators but also it importantly enables point quadrat sampling to be embedded within a regression framework. The methods provided by MacArthur and Horn (1969) and Radtke and Bolstad (2001) can be used to estimate LAI and to obtain a graph of the canopy profile, but they are limited in their ability to incorporate stand-level covariates to improve canopy profile and LAI estimation.

## Regression Analysis of Point Quadrat Data

Radtke and Bolstad (2001) note that the number of sample points needed to obtain accurate estimates using laser point quadrat sampling can be quite high. They observed minimal improvement in LAI estimation over existing LAI estimation methods with a sample size of approximately 1,000 laser points per  $13 \times 13$ -m plot. Fukushima et al. (1998) suggest that adequate foliage profile estimates can be obtained with a minimum of 1,600 sample points distributed uniformly across the stand. These large sample sizes make direct inference of a single stand difficult and make comparison between many stands impractical using traditional point quadrat analysis methods.

Survival-based regression analysis of point quadrat data can address this situation by incorporating stand-specific covariates into the LAI estimation process. This enhanced methodology can lead to improved estimates of LAI and canopy profile, and it allows for quantitative identification of those stand-level variables that affect canopy structure and LAI. Survival-based regression analysis can potentially result in fewer sample points by combining information across plots.

## Study Design

As a simple illustration of the usefulness of the survival-based regression framework for analyzing point quadrat data, we conducted a field study in the counties of Union and Baker in northeast Oregon. Annual precipitation in this region is generally less than 50 cm/year, and the study sites were dominated by dry upland forest communities comprised of ponderosa pine (*Pinus ponderosa*), Douglas-fir (*Pseudotsuga menziesii*), grand fir (*Abies grandis*), and western larch (*Larix occidentalis*), with lesser percentages of lodgepole pine (*Pinus contorta*), Engelmann spruce (*Picea engelmannii*), and subalpine fir (*Abies lasiocarpa*).

We subjectively selected 60 plots across these two counties in order to obtain a wide range of forest types and structures. Within each of the 60 stands, a plot center was randomly selected. At each plot, a series of three sample points was established, each spaced 30-m apart in a north-south orientation. At each point, we tallied trees using a Spiegel Relaskop with a metric basal area factor of  $4 \text{ m}^2 \text{ ha}^{-1}$  ( $17.4 \text{ ft}^2 \text{ ac}^{-1}$ ) and recorded the species name and diameter at breast height (dbh) for each of these trees. To obtain approximately 10–12 tree height measurements per plot, we used a variable basal area factor of either  $4 \text{ m}^2 \text{ ha}^{-1}$  ( $17.4 \text{ ft}^2 \text{ ac}^{-1}$ ),  $9 \text{ m}^2 \text{ ha}^{-1}$  ( $39.2 \text{ ft}^2 \text{ ac}^{-1}$ ), or  $16 \text{ m}^2 \text{ ha}^{-1}$  ( $69.7 \text{ ft}^2 \text{ ac}^{-1}$ ) to identify a subset of trees, from which we obtained height measurements using a clinometer. Based on the dominant and codominant tree species at each plot, we categorized the sites as (1) ponderosa pine, (2) ponderosa mix, (3) mixed conifer, (4) lodgepole pine, or (5) subalpine fir forest types.

Within each stand, we established a 90-m transect, centered at the middle sample point and extending 15 m beyond the northern and southern sample points. A Leica DISTO D8 Laser Rangefinder with tilt sensor was used to measure height-to-first-leaf along this transect. This laser rangefinder has a range of 0.05–200 m with an accuracy of approximately 1 mm. It includes a “long-range mode,” which we used in bright sunlight or for very high canopies. Using this laser rangefinder, a field crew member walked along the transect and took a height-to-canopy measurement every 1 m, with no measurement taken at the 0-m mark. The rangefinder was held at breast height (approximately 1.3 m above the ground), and the tilt sensor was used to ensure that the unit was held vertically. Using the built-in wireless capability of the DISTO D8, the height measurement at each location along the transect was immediately transmitted to a handheld tablet computer being operated by another field crew member.

This procedure resulted in 90 foliage height measurements for each stand. Sky hits (those points that passed through the entire canopy without intercepting any foliage) returned an error and could not be transmitted to the tablet computer. The number of sky hits for each stand was therefore indicated by the number of missing measurements (out of 90) for that stand.

## Statistical Analysis

The Weibull distribution was used to model laser intercept heights. The probability density function for the two-parameter Weibull is

$$f(h|\alpha, \lambda) = \alpha \cdot h^{\alpha-1} \exp(\lambda - \exp(\lambda)h^\alpha) \quad (15)$$

where:  $-\infty < \lambda < \infty; \alpha > 0$ ,

with the corresponding survival function

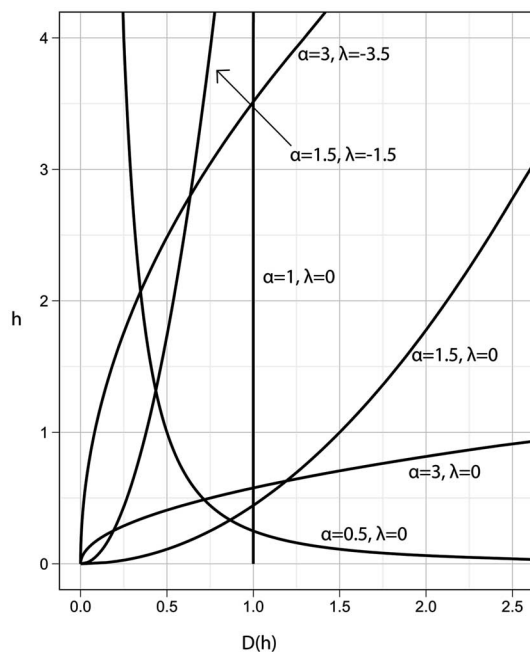
$$S(h|\alpha, \lambda) = \exp(-\exp(\lambda)h^\alpha). \quad (16)$$

The canopy density function is specified by the Weibull hazard function

$$D(h|\alpha, \lambda) = \frac{f(h|\alpha, \lambda)}{S(h|\alpha, \lambda)} = \alpha \cdot h^{\alpha-1} \exp(\lambda). \quad (17)$$

This hazard function provides a range of monotonic increasing, decreasing, and constant canopy density profiles (Figure 1). The possible forms for  $D(h)$  given by the Weibull hazard function are undoubtedly oversimplifications of real-world canopy profiles because of the lack of unimodal or bimodal hazard functions, but the flexibility of these potential canopy density profiles makes the Weibull hazard function ideal for illustrative purposes and for modeling different forest types with disparate canopy structures. In addition, this distribution has previously been used for modeling canopy structure in various forms (Schreuder and Swank 1974, Moril and Hagihara 1991, Yang et al. 1993, Coops et al. 2007), although not via the hazard function. Alternative parametric functions that are commonly used in survival analysis (including the exponential, log-normal, log-logistic, Cauchy, and gamma distributions) were considered but were ultimately not selected owing to various limitations.

In the regression setting, covariates are modeled by the link function  $\lambda = \mathbf{X}_j\boldsymbol{\beta}$ , where  $\mathbf{X}_j = (X_{1j}, X_{2j}, \dots, X_{sj})$  is the covariate vector for the  $j$ th site, and  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_s)$  is the vector of regression coefficients. This gives



**Figure 1.** Various density functions given by the two-parameter Weibull distribution. Although limited to monotonic functions, the Weibull hazard function provides a wide range of canopy profile shapes.

$$D(h|\mathbf{X}_j, \alpha, \beta) = \frac{f(h|\mathbf{X}_j, \alpha, \beta)}{S(h|\mathbf{X}_j, \alpha, \beta)} = \alpha \cdot h^{\alpha-1} \exp(\mathbf{X}_j\boldsymbol{\beta}). \quad (18)$$

An attractive aspect of the Weibull distribution in the survival analysis setting is the intuitive interpretation of the regression coefficients. Each term  $\exp(\beta_k)$  in Equation 18 is equal to a relative difference in the density (or LAI) between two stands for a 1-unit increase in the variable  $X_k$ , holding all other covariates fixed. For example, suppose that two sites have identical covariates, except that the value of the  $k$ th covariate for site 1 is exactly 1 unit greater than that of site 2; that is,  $X_{k1} = X_{k2} + 1$ . Then we have that the canopy density ratio between the two stands is

$$\begin{aligned} \frac{D_B(h|\mathbf{X}_1, \alpha, \beta)}{D_B(h|\mathbf{X}_2, \alpha, \beta)} &= \frac{\alpha \cdot h^{\alpha-1} \exp(\beta_1 \cdot X_{11} + \beta_2 \cdot X_{21} + \dots + \beta_s \cdot X_{s1})}{\alpha \cdot h^{\alpha-1} \exp(\beta_1 \cdot X_{12} + \beta_2 \cdot X_{22} + \dots + \beta_s \cdot X_{s2})} \\ &= \frac{\exp[\beta_k \cdot (X_{k2} + 1)]}{\exp(\beta_k \cdot X_{k2})} \\ &= \exp(\beta_k). \end{aligned} \quad (19)$$

Accordingly,  $\exp(\beta_k)$  represents the ratio of canopy densities between two stands for a 1-unit difference in the  $k$ th covariate. In survival analysis, this is termed the “hazard ratio,” as it represents the ratio between hazards for a 1-unit increase in the corresponding variable. Here, we refer to  $\exp(\beta_k)$  as the “LAI ratio” for the  $k$ th covariate.

### ***Incorporating Sky Hits into the Regression Framework***

In the nonparametric methods proposed by MacArthur and Horn (1969) and Radtke and Bolstad (2001), sky hits do not present any issues. In the parametric framework, however, sky hits propose a unique challenge, as their survival times are infinite. These points cannot be trivially excluded from the analysis, as they are informative with respect to total leaf area. They also cannot be included as censored observations, as such observations are still assumed to have finite failure times.

To avoid the problem of sky hits in the regression setting, we can conceptually partition the total stand area,  $A$ , into the area beneath foliage,  $A_B$ , and the area not beneath foliage,  $A_{\bar{B}}$ . By ignoring the difference between foliage and woody canopy material, this latter area includes all of the locations in the stand that would result in sky hits. Implicit in this partitioning is the assumption that foliage movement is minimal and that the gap fraction is roughly constant over short periods of time. From Equation 2, we see that the canopy density function can be written as the derivative of the LAI function. That is,

$$D(h) = \lim_{\delta \rightarrow 0} \frac{\text{LAI}(0, h - \delta) - \text{LAI}(0, h)}{\delta}. \quad (20)$$

By definition,  $\text{LAI}(0, h) = \text{LA}(h)/A$ ; where  $\text{LA}(h)$  is the total leaf area across the stand up to height  $h$ . Because the portion

of the stand that is not beneath foliage contains zero leaf area at all heights, it follows that  $LA(h) = LA_B(h)$ , where  $LA_B(h)$  is the leaf area up to height  $h$  for the portion of the stand beneath foliage.

If  $D_B(h)$  and  $LAI_B(0, h)$  denote the foliage density and LAI up to height  $h$ , respectively, for the portion of the stand beneath foliage, then it follows that

$$\begin{aligned}
 D(h) &= \lim_{\delta \rightarrow 0} \frac{LAI(0, h + \delta) - LAI(0, h)}{\delta} \\
 &= \lim_{\delta \rightarrow 0} \frac{LA(h + \delta)/A - LA(h)/A}{\delta} \\
 &= \frac{1}{A} \lim_{\delta \rightarrow 0} \frac{LA(h + \delta) - LA(h)}{\delta} \\
 &= \frac{A_B}{A} \lim_{\delta \rightarrow 0} \frac{LA(h + \delta)/A_B - LA(h)/A_B}{\delta} \\
 &= \frac{A_B}{A} \lim_{\delta \rightarrow 0} \frac{LA_B(h + \delta)/A_B - LA_B(h)/A_B}{\delta} \\
 &= \frac{A_B}{A} \lim_{\delta \rightarrow 0} \frac{LAI_B(0, h + \delta) - LAI_B(0, h)}{\delta} \\
 &= \frac{A_B}{A} D_B(h) \\
 &= (1 - g) \cdot D_B(h). \tag{21}
 \end{aligned}$$

Intuitively, this result is equivalent to an area-weighted canopy density function

$$\begin{aligned}
 D(h) &= \frac{A_B}{A} \cdot D_B(h) + \frac{A_{\bar{B}}}{A} \cdot D_{\bar{B}}(h) \\
 &= (1 - g) \cdot D_B(h) + g \cdot 0 \\
 &= (1 - g) \cdot D_B(h), \tag{22}
 \end{aligned}$$

where  $D_{\bar{B}}(h)$  is the density function at height  $h$  for the portion of the stand not beneath foliage (and is trivially equal to 0 for all values of  $h$ ). This area-weighted canopy density function is identical to a hazard mixture model in industrial and technological survival analysis settings (Abdelfattah and Mustafa 2012), in which a device can fail for one of two reasons, but it cannot be at risk of failing for both reasons simultaneously.

An estimate of the overall density function can be obtained by an estimate of the gap fraction and an estimate of the density function using only those points that intercepted foliage

$$\hat{D}(h) = (1 - \hat{g}) \cdot \hat{D}_B(h). \tag{23}$$

An issue with estimating this model using a parametric hazard function is that all parametric hazard functions satisfy the condition that their integral on  $(0, \infty)$  is equal to infinity. To alleviate this problem, we can calculate the total

leaf area of the stand by integrating the density function only up to the upper limit of the canopy

$$LAI = \int_0^{h_m} D(h) dh = (1 - g) \int_0^{h_m} D_B(h) dh < \infty, \tag{24}$$

which is equivalent to truncating the distribution of canopy foliage at the max tree height,  $h_m$ .

An important implication of this result is that estimation of the survival regression parameters can be conducted using only the intercepted laser hits, thus eliminating the difficulty of incorporating infinite survival times into the model. Under this framework, estimation of the overall parametric density equation is a four-step process: (1) estimate  $D_B(h)$  using only the laser hits that intercepted foliage, (2) estimate  $g$  using the proportion of sky hits, (3) estimate  $D(h)$  by taking the product of  $D_B(h)$  and  $(1 - g)$ , and (4) truncate  $D(h)$  at the maximum canopy height.

### Canopy Profiles and LAI Estimates

To allow for greater flexibility in the shape of the canopy density function between forest types, the Weibull regression model given in Equation 18 was fit separately for ponderosa pine, ponderosa pine mix, and mixed-conifer forest types. Using the above framework, we estimated the coefficients using only those laser shots that intercepted foliage. In addition, a value of 1.3 m was added to each laser intercept height, as the laser rangefinder was held approximately 1.3 m above the ground.

The covariates included in the regression models were basal area per hectare (BA), trees per hectare (TPH), stand density index (SDI), and indicator variables denoting the ownership type for each plot (private nonindustrial, private industrial, and USDA Forest Service [USFS] land). SDI was calculated using an additive version of Reineke's SDI given by Long and Daniel (1990), and it was first calculated for each plot and then averaged across the three plots at each site. In addition, as specified in Gove (2003), we fit a two-parameter size-biased Weibull model to the dbh measurements for each plot to estimate the complete diameter distribution of trees in that stand. The shape and scale parameters from these size-biased Weibull diameter models were included as additional covariates. Quadratic mean diameter (QMD) was originally included but was removed because of high collinearity with the size-biased Weibull diameter distribution parameters. Site-specific differences were originally modeled through a random-effects term (known as a "frailty" model in survival analysis). These results did not appreciably differ from the fixed-effects model, so this random-effects term was removed for parsimony and to facilitate site-specific estimation of the canopy profiles and LAI values.

For each site, we estimated both parametric and nonparametric canopy profiles. For the parametric canopy profiles, the coefficients in Equation 18 were estimated using maximum likelihood estimation. For each of the three forest types, the likelihood function for the vector  $\mathbf{h} = (h_1, \dots, h_n)$  of leaf intercept heights is given by

$$L(\alpha, \beta | \mathbf{h}, \tilde{\mathbf{X}}) = \prod_{i=1}^n [f(h_i | \mathbf{X}_i, \alpha, \beta)^{v_i} \cdot S(h_i | \mathbf{X}_i, \alpha, \beta)^{1-v_i}] \quad (25)$$

where  $v_i$  is a vector of indicator variables denoting if the  $i$ th laser shot was censored ( $v_i = 0$ ) or observed ( $v_i = 1$ ), and  $\tilde{\mathbf{X}}$  is the matrix of the site-specific covariate vectors corresponding to each of the  $n$  leaf intercept height measurements. Although the number of measured leaf intercept heights,  $n$ , is a random variable, it is treated as a fixed constant to simplify estimation.

Substituting Equations 15 and 16 into Equation 25 (using the link  $\lambda = \mathbf{X}_i'\beta$ ) and rearranging terms, it follows that the log-likelihood function is

$$l(\alpha, \beta | \mathbf{h}, \tilde{\mathbf{X}}) = \sum_{i=1}^n v_i (\log(\alpha) + (\alpha - 1) \log(h_i) + \mathbf{X}_i'\beta) - \exp(\mathbf{X}_i'\beta) h_i^\alpha \quad (26)$$

In this study, we did not incorporate censoring, and so  $v_i = 1$  for all leaf intercept measurements. Parameter estimation of the regression coefficients in Equation 26 was conducted using a Newton-Raphson iteratively reweighted least-squares algorithm embedded within the *survival* package in R (R Foundation for Statistical Computing, Vienna, Austria, 2010). See Therneau (1999) for an explanation of the specific estimation algorithm used by this statistical package and Green (1984) for an overview of iteratively reweighted least-squares maximum likelihood estimation.

After estimation of regression coefficients, the site-specific density functions from Equation 18 were each multiplied by the corresponding empirical gap fraction for that site (the proportion of sky hits at that site), resulting in a smooth estimate of the canopy profile as given in Equation 23. The graphs were constructed by plotting  $\hat{D}(h)$  against  $h$  and truncating the curve at the maximum tree height for each site (using the maximum of the recorded tree heights and laser intercept heights for that site). Parametric regression-based LAI values were obtained by integrating  $\hat{D}(h)$  up to the maximum measured tree height, as indicated in Equation 24. We also considered using the average of the recorded tree heights, but the results did not differ appreciably from using the maximum recorded tree height.

Nonparametric canopy profiles were obtained by dividing the canopy height into 1-m intervals and using Equation 6 to calculate the Radtke-Bolstad LAI estimate within each

of these intervals for each site (analogous to the Nelson-Aalen estimate). These interval-specific LAI estimates were then plotted using a stepwise linear function. Nonparametric LAI values were estimated using the Kaplan-Meier product-limit estimator (equivalent to the MacArthur-Horn estimator as given in Equation 4). As with the parametric canopy profiles, the maximum canopy height used to calculate the parametric LAI was estimated as the maximum of the recorded tree heights or laser intercept heights for each plot. The *muhaaz* package in R was used for nonparametric estimation of the LAI estimates and canopy profiles.

LAI ratios for each forest type and covariate were calculated by exponentiating the estimated regression coefficients for each of the three forest-type survival models. For each of these LAI ratios, the asymmetric 95% confidence intervals (CIs) were obtained by exponentiating the 95% CI of the corresponding regression coefficient (assuming a Student's  $t$ -distribution). CIs for the regression-based LAI estimates and the canopy profiles are not given, because additional uncertainty is present in the estimate of the gap fraction and the maximum canopy height. These point estimates should therefore be interpreted with caution, and they are reported to allow for comparisons between the nonparametric and parametric approaches.

## Results

Of the 60 sites, we classified 22 as mixed conifer, 19 as ponderosa pine, 14 as ponderosa mix, 4 as subalpine fir, and 1 as lodgepole pine. The subalpine fir and lodgepole pine sites were omitted due to low numbers, resulting in 55 sites used in the analysis. Of these remaining sites, 18 were on USFS land, 16 were on private nonindustrial land, and 21 were on private industrial land.

The sites were primarily open, with gap fractions ranging from approximately 60–70% (as estimated by the fraction of sky hits). The sites in each of the three remaining forest types were generally comparable in QMD, height, and live crown ratio (Table 1). The mixed-conifer stands showed fewer signs of active management and/or timber harvests. Accordingly, these stands had greater crown closure and denser foliage throughout the canopy strata. These mixed conifer sites had slightly more BA than ponderosa and ponderosa mix stands (29.5 m<sup>2</sup> ha<sup>-1</sup> compared with 21.8 and 21.0 m<sup>2</sup> ha<sup>-1</sup>, respectively; Table 1). TPH and SDI were also markedly larger in mixed conifer stands.

Regression-based canopy profiles for all sites are given

**Table 1. Summary statistics of the 55\* forest plots.**

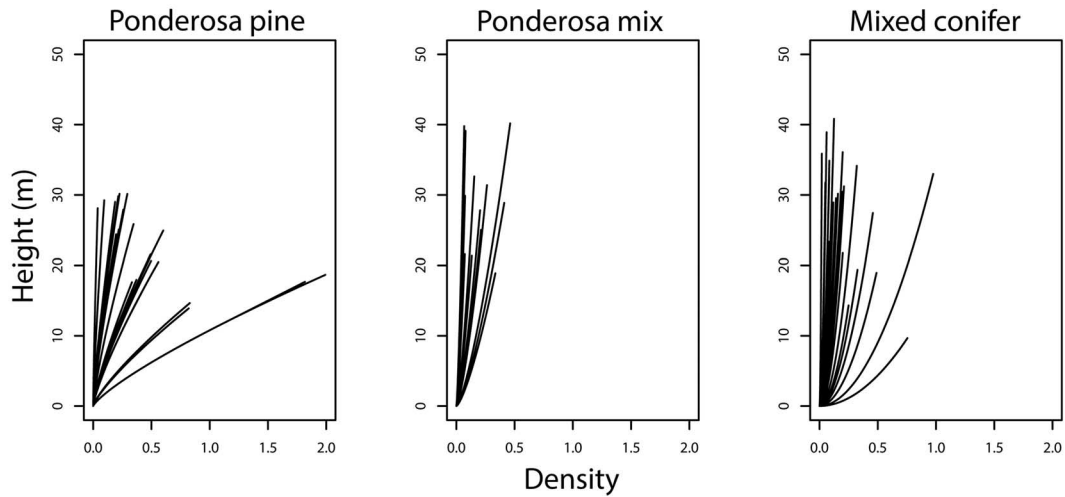
Variable	Mixed conifer ( $n = 22$ )	Ponderosa pine ( $n = 19$ )	Ponderosa mix ( $n = 14$ )
BA (m <sup>2</sup> ha <sup>-1</sup> )	29.5 ± 12.6	21.8 ± 12.6	21.0 ± 9.1
QMD (m)	0.26 ± 0.10	0.30 ± 0.09	0.32 ± 0.07
TPH	1,036.9 ± 1,610.2	353.2 ± 251.1	313.6 ± 214.7
SDI	52.6 ± 58.4	30.6 ± 20.6	41.1 ± 21.7
Height (m)	19.8 ± 6.4	17.7 ± 5.1	21.8 ± 4.9
Live crown ratio (%)	28 ± 13	30 ± 11	28 ± 8
Gap fraction (%)†	60 ± 19	68 ± 13	68 ± 12

Data are means ± SD.

\* Lodgepole pine and subalpine fir sites ( $n = 5$ ) are omitted.

† Estimated using the proportion of laser sky hits.



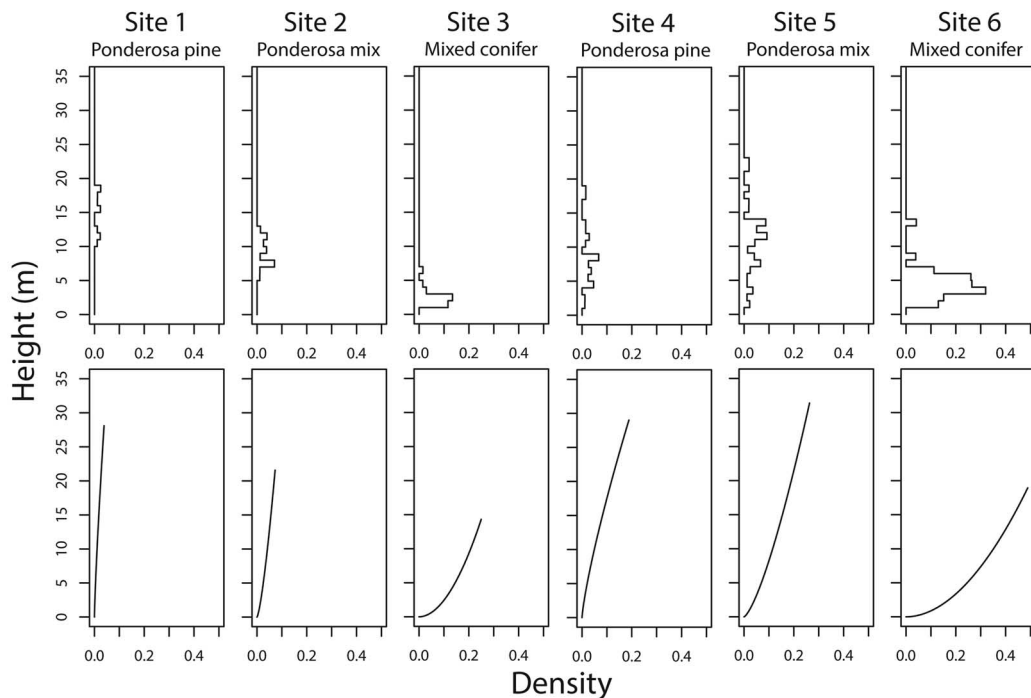


**Figure 2.** The estimated canopy profiles by forest type for all sites using the Weibull regression model. The maximum canopy profile for each site is equal to the maximum measured tree height (or laser intercept height) at that site. The canopy area contained within each profile is equal to the parametric LAI estimate for that site.

in Figure 2. The ponderosa pine sites were generally the shortest of the three forest types, with minimal foliage density below 10 m and foliage density that increased relatively quickly near the upper portion of the canopy. In contrast, the ponderosa mix sites showed more uniform foliage density throughout all strata. The mixed-conifer stands showed the most even distribution of foliage throughout the canopy, with the majority of stands having very little change in foliage density from 10 m in height to the top of the canopy.

Because of the small number of laser intercepts for each stand (approximately 30 non-sky hits for each site), very few laser shots penetrated into the upper portions of the canopy. Accordingly, the nonparametric profiles drastically

underestimate canopy density, with no foliage or woody material estimated above 25 m for any of the stands (Figure 3). Conversely, the parametric regression model interpolated the canopy density profiles to the upper portions of the canopy. For stands with dense foliage close to the ground (e.g., sites 3 and 6 of Figure 3), the parametric and nonparametric profiles show the largest discrepancies. For stands with less dense foliage (e.g., sites 1 and 2), the parametric and nonparametric profiles are much more similar, because the laser shots were able to penetrate farther into the canopy. This improvement by the parametric canopy profiles in dense stands highlights the benefit of a parametric approach: under the assumption that the underlying probability model is approximately correct, height



**Figure 3.** The estimated canopy profiles for six different sites. The top row gives the nonparametric estimates and the bottom row gives the corresponding parametric estimates.



measurements at the upper portions of the canopy are not needed to infer the density at that height.

The parametric LAI estimates reflect these qualitative differences in canopy profile form. The mixed conifer sites had the largest mean LAI estimate, with a value of 4.4 (95% CI = 1.7–7.0). The ponderosa pine sites had a substantially higher mean LAI estimate than the ponderosa mix sites, with a value of 4.1 (95% CI = 2.6–5.6) compared with 2.4 (95% CI = 1.4–3.4).

The MacArthur-Horn estimates of LAI were markedly less than the corresponding parametric values (Figure 4). The empirical 5 and 95% quantiles for the regression-based LAI estimates and MacArthur-Horn LAI estimates were (0.9, 9.0) and (0.1, 1.0), respectively, showing large differences in the LAI estimates between the two methods.

The estimated LAI ratios from the three Weibull regression models are given in Table 2. These ratios correspond to the relative difference in LAI between two stands that differ only by one unit of the reference variable and have all remaining covariates identical. The results show that private industrial corresponds to significantly higher LAI for all three forest types than identical sites on USDA Forest Service land, with LAI ratios of 1.63 (95% CI = 1.23–2.15), 1.63 (95% CI = 1.20–2.22), and 2.38 (95% CI = 1.92–2.96) for ponderosa, ponderosa mix, and mixed-conifer forest types, respectively. Conversely, a site on private property is estimated to have a reduction in LAI by a factor of 0.70 (95% CI = 0.55–0.90) compared with a similar site on USDA Forest Service property.

Across all three sites, an increase in basal area corresponded to a decrease in canopy density (potentially due to increased competition and decreased growing space). Specifically, a  $5\text{-m}^2\text{ ha}^{-1}$  decrease in BA resulted in a reduction in LAI by factors of 0.81, 0.89, and 0.80 for ponderosa, ponderosa mix, and mixed-conifer sites, respectively. Although this result may seem counterintuitive, it is important

to realize that this relationship is the result of holding all other variables constant (including the gap fraction). Neither TPH nor SDI showed any obvious trends. An increase in TPH led to an increase in density for ponderosa and mixed-conifer sites, but it led to a decrease in density for ponderosa mix sites. Conversely, an increase in SDI corresponded to an increase in LAI among ponderosa mix and mixed-conifer sites, but a decrease in LAI for ponderosa sites.

## Discussion

The primary objectives of this study were to outline the relationship between survival analysis and laser point quadrat sampling and to provide a simple example that illustrates the utility of this relationship. In this study we used only 90 sample points per stand, with approximately 35% of the laser shots intercepting foliage. For comparison, Radtke and Bolstad (2001) reported mixed results with approximately 1,000 sample points per stand, and Fukushima et al. (1998) recommended approximately 1,600 sample points per stand using the traditional MacArthur-Horn estimators. By using a parametric approach and incorporating covariates into a regression framework, we were able to obtain smooth canopy profiles and LAI estimates with a fraction of the sample points used in previous studies. Direct validation of these parametric estimates was beyond the scope of this study, although the results allow for qualitative comparisons between the nonparametric and survival-based approaches to analyzing laser point quadrat data.

To improve on the applicability of combining survival analysis and point quadrat sampling, additional research is needed in five distinct areas: (1) appropriate selection of the density (hazard) function, (2) simpler methods for incorporating sky hits, (3) techniques for addressing multiple sources of error in the estimation process, (4) tools for distinguishing LAI from plant area index, and (5) techniques for adjusting for clumping and interdependence between sample points.

Although we selected the Weibull hazard function for its simplicity, flexibility, and interpretability, it is probably ill-suited for those situations in which highly detailed canopy profile estimates are needed. The parametric canopy profiles shown in this article allow for general, qualitative comparisons between forest canopies, but the available canopy profile forms were limited by the monotonic nature of the Weibull hazard function. Overall, selection of an appropriate hazard function for laser survival times is difficult. Distributions such as the log-normal, log-logistic, and Cauchy have unimodal hazard functions, but the maximum values of these hazard functions (and accordingly, the height of the maximum canopy density) typically occur much closer to the ground than would be seen in most forest stands. Additional distributions common to survival analysis were also investigated for use in this study (including the exponential distribution, the normal distribution, the *t*-distribution, the gamma distribution, the extreme-value distribution, and the uniform distribution) but were likewise not used either because the corresponding hazard profiles did not reflect standard canopy profiles or because the hazard

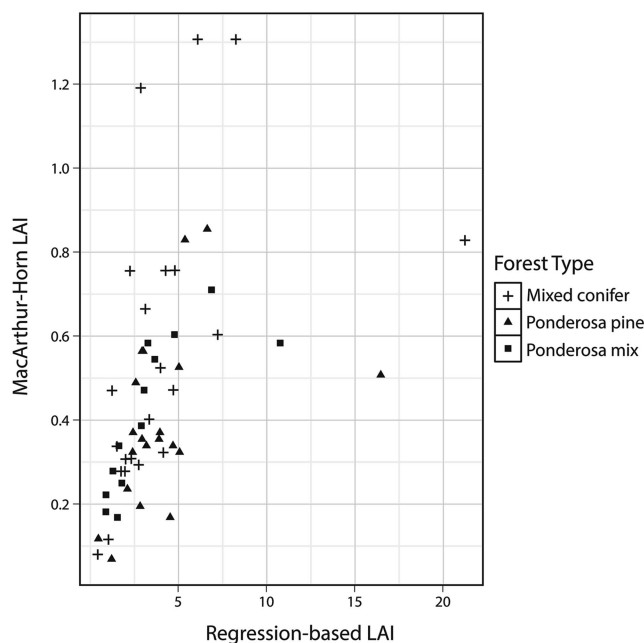


Figure 4. The estimated MacArthur-Horn LAI values and the regression-based LAI values for each site by forest type.

**Table 2. Results from the Weibull regression models for ponderosa pine, ponderosa pine mix, and mixed-conifer forest types.**

Covariate	LAI ratio*	95% CI	P value
Ponderosa pine			
Private property (versus USFS)	0.70	0.55–0.90	0.005
Industrial property (versus USFS)	1.63	1.23–2.15	0.001
BA†	0.81	0.66–1.00	0.053
TPH‡	1.39	1.12–1.72	0.003
SDI§	0.77	0.61–0.99	0.039
Diameter: Weibull shape	1.17	1.09–1.24	<0.001
Diameter: Weibull scale	4.40	0.03–672.9	0.563
Ponderosa pine mix			
Private property (versus USFS)	1.29	0.70–2.37	0.418
Industrial property (versus USFS)	1.63	1.20–2.22	0.002
BA†	0.89	0.78–1.02	0.099
TPH‡	0.92	0.84–1.00	0.056
SDI§	1.15	1.04–1.28	0.007
Diameter: Weibull shape	1.18	1.06–1.33	0.004
Diameter: Weibull scale	0.001	0.0002–0.01	<0.001
Mixed conifer			
Private property (versus USFS)	1.81	1.34–2.44	<0.001
Industrial property (versus USFS)	2.38	1.92–2.96	<0.001
BA†	0.80	0.76–0.85	<0.001
TPH‡	1.02	1.01–1.02	<0.001
SDI§	1.10	1.06–1.15	<0.001
Diameter: Weibull shape	0.83	0.72–0.96	0.015
Diameter: Weibull scale	0.02	0.003–0.19	<0.001

\* Equals the relative change in LAI and canopy density, holding all other covariates and the gap fraction fixed.

† Corresponds to a 5 m<sup>2</sup> ha<sup>-1</sup> increase.

‡ Corresponds to a 100 TPH increase.

§ Corresponds to a 10-unit increase.

function did not provide the desired flexibility in canopy profiles. The suitability of additional distributions not typically used in survival analysis will need to be investigated on a one-by-one basis, as the shape of the hazard function is often not known or is rarely reported.

To achieve more realistic canopy profiles using survival analysis, more complex modeling methods or distributions may be needed. Finite mixture models (Schlattmann 2009) would allow for the canopy density function to be a modeled using a combination of simple hazard functions. This could be used in forests with bimodal or nonstandard canopy structures. Bayesian methods (both parametric and nonparametric) would also probably improve canopy profile estimation by allowing for the specification of a baseline canopy density function that is then updated to reflect the laser intercept heights and covariates for each stand. Appropriate selection of this baseline (prior) distribution would allow for more realistic canopy profiles that approximate known canopy structures and borrow results from past research.

An additional possibility for improving hazard estimation is to take a nonparametric or semiparametric approach. The Weibull distribution used here is an example of a proportional hazards models, because the hazard functions differ only by a constant of proportionality given by their covariates. More general proportional hazards models make no assumptions about the underlying baseline hazard. If the assumption of proportional hazards holds, then it is possible to estimate the relative impact of each covariate without specifying the failure-time distribution (Cox 1972), thus limiting the effect of model misspecification. We considered various nonparametric approaches to modeling canopy

profiles, such as using a piecewise constant hazard model, but such models offered no improvement over the traditional nonparametric MacArthur-Horn approach with regards to sample size, accuracy, or resulting canopy profile shapes.

More parsimonious regression approaches than the one used in this study may also be possible. Because of sky hits, we used a two-step procedure to estimate the density function. A drawback of the parametric estimation procedure that we selected is that error is present in three places: the estimate of the gap fraction, the estimate of  $\hat{D}(h)$ , and the estimate of the upper portion of the canopy profile. For this reason, statistically valid 95% CIs and SEs were not calculable for the parametric canopy profiles or LAI estimates. Although this is undoubtedly a limitation of this approach, we intend the methods presented here to serve as a useful starting point for future research in this area. Ideally, it would be desirable to develop a methodology that either does not require data partitioning or allows for simultaneous estimation of the density parameters and the gap fraction parameters. In hope of avoiding this two-step procedure, we investigated the use of infinite-valued distributions, mixture distribution, and cure fraction models, where a subset of patients are “cured” and no longer at risk of failure (De Angelis et al. 1999, Lambert et al. 2007). All of these models allow for a partitioning of the data into two general cohorts (i.e., those points beneath foliage and not beneath foliage), but they all invoke additional constraints that are not immediately relevant in the point quadrat setting. Further research into the use of these distributions in laser point quadrat sampling is needed.

As discussed, our use of the term “leaf area index” throughout this article refers to the “plant area index” because both foliage and woody material returned laser distance readings. Many laser rangefinders (including the Leica DISTO used in this study) include a digital display that allows the user to identify the exact object that is reflecting the laser shot. By recording this additional information at each laser hit, two existing survival analysis methods can be used to distinguish LAI from plant area index. The first method involves modeling the woody laser hits as censored values, because these shots were intercepted (censored) by woody material before they could be intercepted by foliage. Including these censored values into a regression model for failure time data is simple, because most statistical analysis packages for survival analysis are specifically designed to incorporate censored values. A second method for incorporating woody material is via a “competing risks model” (Kalbfleisch and Prentice 2002). In these models, multiple sources of failure are accounted for, and hazard function estimation is conducted for each source. In the point quadrat setting, this would allow for the estimation of two separate density functions: one corresponding to LAI and one corresponding to wood area index. Methods for implementing competing risk models are likewise available in many statistical software packages.

Last, clumping of foliage along with the interdependence of foliage remains an issue with point quadrat sampling. The presence of strong clumping (particularly in coniferous stands) can make it impossible for the laser to reach some foliage that is totally obscured by other foliage. This limitation has been acknowledged and well-documented elsewhere (see Radtke and Bolstad 2001), although there are no well-established methods of adjusting for clumping in the laser point quadrat setting.

## Conclusions

Survival analysis methods provide a promising set of statistical tools for analyzing point quadrat data. Although additional research is needed to explore the effectiveness and limitations of this method, survival-based point quadrat sampling is a useful technique for quickly obtaining LAI estimates, modeling canopy profiles, and identifying stand-level covariates that affect canopy density. The flexibility and simplicity of survival-based regression models make them ideal for identifying stands with poor canopy structures, for conducting widespread forest health studies or for assessing wildlife habitat. Although other optical sampling methods such as LiDAR offer much higher spatial resolution, the advantage of laser point quadrat sampling is that rangefinders are inexpensive, easy to operate, and portable, and little is needed in terms of operational infrastructure. In addition, the laser point quadrat sampling method we used took less than 5 minutes to complete per plot, and the ancillary regression covariates we included are typically already gathered in most forest inventory procedures. Overall, the relationship between point quadrat sampling and survival analysis increases the efficiency and functionality of point quadrat sampling, resulting in a useful tool for

quickly and accurately assessing the structure of forest canopies.

## Literature Cited

- AALLEN, O.O. 1978. Nonparametric inference for a family of counting processes. *Ann. Statist.* 6(4):701–726.
- ABDELFATTAH, A., AND A. MUSTAFA. 2012. Reliability equivalence factors for some systems with mixture Weibull failure rates. *Am. J. Math. Stat.* 2(1):6–12.
- ABER, J.D. 1979. Foliage-height profiles and succession in northern hardwood forests. *Ecology* 60(1):18–23.
- BALDOCCHI, D.D., AND P.C. HARLEY. 2006. Scaling carbon dioxide and water vapour exchange from leaf to canopy in a deciduous forest. II. Model testing and application. *Plant Cell Environ.* 18(10):1157–1173.
- BALDOCCHI, D.D., K.B. WILSON, AND L. GU. 2002. How the environment, canopy structure and canopy physiological functioning influence carbon, water and energy fluxes of a temperate broad-leaved deciduous forest—an assessment with the biophysical model CANOAK. *Tree Physiol.* 22:1065–1077.
- BOLSTAD, P.V., AND S.T. GOWER. 1990. Estimation of leaf area index in fourteen southern Wisconsin forest stands using a portable radiometer. *Tree Physiol.* 7:115–124.
- CHEN, J.M., AND T.A. BLACK. 1992a. Defining leaf area index for non-flat leaves. *Plant Cell Environ.* 15:421–429.
- CHEN, J.M., AND T.A. BLACK. 1992b. Foliage area and architecture of plant canopies from sunfleck size distributions. *Agr. For. Meteorol.* 60:249–266.
- CHEN, J.M., AND J. CIHLAR. 1995. Plant canopy gap-size analysis theory for improving optical measurements of leaf-area index. *Appl. Opt.* 34(27):6211–6222.
- COOPS, N.C., T. HILKER, M.A. WULDER, B. ST-ONGE, G. NEWNHAM, A. SIGGINS, AND J.A.T. TROFYMOW. 2007. Estimating canopy structure of Douglas-fir forest stands from discrete-return LiDAR. *Trees* 21(3):295–310.
- CÔTÉ, J.F., R.A. FOURNIER, AND R. EGLI. 2011. An architectural model of trees to estimate forest structural attributes using terrestrial LiDAR. *Environ. Model. Softw.* 26:761–777.
- COX, D.R. 1972. Regression models and life-tables. *J. R. Stat. Soc. Ser. B.* 34(2):187–220.
- DE ANGELIS, R., R. CAPOCACCIA, T. HAKULINEN, B. SODERMAN, AND A. VERDECCHIA. 1999. Mixture models for cancer survival analysis: Application to population-based data with covariates. *Stat. Med.* 18:441–454.
- DE MAYNADIER, P.G., AND M.L. HUNTER JR. 1999. Forest canopy closure and juvenile emigration by pool-breeding amphibians in Maine. *J. Wildl. Manage.* 63(2):441–450.
- DENISON, R.F. 1997. Minimizing errors in LAI estimates from laser-probe inclined-point quadrats. *Field Crops Res.* 51: 231–240.
- FORD, E.D., AND J.D. DEANS. 1978. The effects of canopy structure on stemflow, throughfall and interception loss in a young Sitka spruce plantation. *J. Appl. Ecol.* 15(3):905–917.
- FUJIMORI, T. 1971. Analysis of forest canopy on the basis of a Tsuga heterophylla stand. *Jpn. J. Ecol.* 59(134–139).
- FUKUSHIMA, Y., T. HIURA, AND S.-I. TANABE. 1998. Accuracy of the MacArthur-Horn method for estimating a foliage profile. *Agr. For. Meteorol.* 92:203–210.
- GOVE, J.H. 2003. Moment and maximum likelihood estimators for Weibull distributions under length- and area-biased sampling. *Environ. Econ. Stat.* 10:455–467.
- GREEN, P.J. 1984. Iteratively reweighted least squares for maximum likelihood estimation and some robust and resistant alternatives. *J. R. Stat. Assoc. B* 46(2):149–192.
- GROENEVELD, D. 1997. Vertical point quadrat sampling and an



- extinction factor to calculate leaf area index. *J. Arid Environ.* 36:475–485.
- HAYES, J.P., S.S. CHAN, W.H. EMMINGHAM, J.C. TAPPEINER, L.D. KELLOGG, AND J.D. BAILEY. 1997. Wildlife response to thinning young forests in the Pacific Northwest. *J. For.* 95(8):28–33.
- HOLLINGER, D.Y. 1989. Canopy organization and foliage photosynthetic montane capacity in a broad-leaved evergreen montane forest. *Funct. Ecol.* 3(1):53–62.
- JENNINGS, S.B., N.D. BROWN, AND D. SHEIL. 1999. Assessing forest canopies and understory illumination: Canopy closure, canopy cover and other measures. *For.* 72(1):59–73.
- JONCKHEERE, I., S. FLECK, K. NACKAERTS, B. MUYS, P. COPPIN, M. WEISS, AND F. BARET. 2004. Review of methods for in situ leaf area index determination: Part I. Theories, sensors and hemispherical photography. *Agr. For. Meteorol.* 121:19–35.
- KALBFLEISCH, J.D., AND R.L. PRENTICE. 2002. *The statistical analysis of failure time data*, 2nd ed. John Wiley and Sons, Hoboken, NJ. 462 p.
- KAPLAN, E.L., AND P. MEIER. 1958. Nonparametric estimation from incomplete observations. *J. Am. Stat. Assoc.* 53:457–481.
- LAMBERT, P.C., J.R. THOMPSON, C.L. WESTON, AND P.W. DICKMAN. 2007. Estimating and modeling the cure fraction in population-based cancer survival analysis. *Biostatistics* 8(3):576–594.
- LEVY, E.B., AND E.A. MADDEN. 1933. The point method of pasture analysis. *N. Z. J. Agr.* 46:267–279.
- LONG, J.N., AND T.W. DANIEL. 1990. Assessment of growing stock in uneven-aged stands. *West. J. Appl. For.* 5(3):93–96.
- LOVELL, J.L., D.L.B. JUPP, D.S. CULVENOR, AND N.C. COOPS. 2003. Using airborne and ground-based ranging LiDAR to measure canopy structure in Australian forests. *Can. J. Remote Sens.* 29(5):607–622.
- MACARTHUR, R.H., AND H.S. HORN. 1969. Foliage profile by vertical measurements. *Ecology* 50(5):802–804.
- MAGUIRE, D.A., AND A. KANASKIE. 2002. The ratio of live crown length to sapwood area as a measure of crown sparseness. *For. Sci.* 48(1):93–100.
- MARTENS, S.N., D.D. BRESHEARS, AND C.W. MEYER. 2000. Spatial distributions of understory light along the grassland/forest continuum: Effects of cover, height, and spatial pattern of tree canopies. *Ecol. Model.* 126:79–93.
- MITCHELL, R.G., R.H. WARING, AND G.B. PITMAN. 1983. Thinning lodgepole pine increases tree vigor and resistance to mountain pine beetle. *For. Sci.* 29(1):204–211.
- MORIL, S., AND A. HAGIHARA. 1991. Crown profile of foliage area characterized with the Weibull distribution in a hinoki (*Chamaecyparis obtusa*) stand. *Trees* 5:149–152.
- NELSON, W. 1972. Theory and applications of hazard plotting for censored failure data. *Technometrics* 14(4):945–966.
- OVINGTON, J.D. 1963. Flower and seed production. A source of error in estimating woodland production, energy flow and mineral cycling. *Oikos* 14(2):148–153.
- PARKER, G.G., D.J. HARDING, AND M.L. BERGER. 2004. A portable LiDAR system for rapid determination of forest canopy structure. *J. Appl. Ecol.* 41:755–767.
- RADTKE, P.J., AND P.V. BOLSTAD. 2001. Laser point-quadrat sampling for estimating foliage-height profiles in broad-leaved forests. *Can. J. For. Res.* 31:410–418.
- RICH, P.M. 1990. Characterizing plant canopies with hemispherical photographs. *Remote Sens. Rev.* 5:13–29.
- ROSS, J. 1981. *Radiation regime and architecture of plant stands*. Springer-Verlag (Kluwer), Hague, The Netherlands. 420 p.
- SCHLATTMANN, P. 2009. *Medical applications of finite mixture models*. Springer, Berlin, Germany. 256 p.
- SCHREUDER, H.T., AND W.T. SWANK. 1974. Coniferous stands characterized with the Weibull distribution. *Can. J. For. Res.* 4:518–523.
- SEIDEL, D., S. FLECK, AND C. LEUSCHNER. 2012. Analyzing forest canopies with ground-based laser scanning: A comparison with hemispherical photography. *Agr. For. Meteorol.* 154–155:1–8.
- SEIDEL, D., S. FLECK, C. LEUSCHNER, AND T. HAMMETT. 2011. Review of ground-based methods to measure the distribution of biomass in forest canopies. *Ann. For. Sci.* 68(2):225–244.
- TABLEMAN, M., J.S. KIM, AND S. PORTNOY. 2004. *Survival analysis using S: Analysis of time-to-event data*. Chapman and Hall/CRC, Boca Raton, FL. 280 p.
- THERNEAU, T.M. 1999. *A package for survival analysis in S*. Mayo Clinic Foundation, Minneapolis, MN, 1999, p. 70–73.
- VANDERBILT, V.C., M.E. BAUER, AND L.F. SILVA. 1979. Prediction of solar irradiance distribution in a wheat canopy using a laser technique. *Agr. Meteorol.* 20:147–160.
- WARING, R.H. 1983. Estimating forest growth and efficiency in relation to canopy leaf area. *Adv. Ecol. Res.* 13:327–354.
- WARING, R.H., K. NEWMAN, AND J. BELL. 1981. Efficiency of tree crowns and stemwood production at different canopy leaf densities. *For.* 54(2):129–137.
- WATSON, D.J. 1947. Comparative physiological studies on the growth of field crops: I. Variation in net assimilation rate and leaf area between species and varieties, and within and between years. *Ann. Bot.* 11:41–67.
- WEISS, M. 2004. Review of methods for in situ leaf area index (LAI) determination. Part II. Estimation of LAI, errors and sampling. *Agr. For. Meteorol.* 121(1–2):37–53.
- WELLES, J.M., AND S. COHEN. 1996. Canopy structure measurement by gap fraction analysis using commercial instrumentation. *J. Exp. Bot.* 47(302):1335–1342.
- WILSON, J.W. 1960. Inclined point quadrats. *New Phytologist* 59(1):1–8.
- WILSON, J.W. 1963. Estimation of foliage denseness and foliage angle by inclined point quadrats. *Aust. J. Bot.* 11:95–105.
- WILSON, J.W. 1965. Stand structure and light penetration. I. Analysis by point quadrats. *J. Appl. Ecol.* 2(2):383–390.
- WILSON, R.F., H. MARSH, AND J. WINTER. 2007. Importance of canopy connectivity for home range and movements of the rainforest arboreal ringtail possum (*Hemibelideus lemuroides*). *Wildl. Res.* 34:177–184.
- YANG, X., D.R. MILLER, AND M.E. MONTGOMERY. 1993. Vertical distributions of canopy foliage and biologically active radiation in a defoliated/refoliated hardwood forest. *Agr. For. Meteorol.* 67:129–146.
- ZHAO, F., X. YANG, M.A. SCHULL, M.O. ROMÁÉN-COLÓN, T. YAO, Z. WANG, Q. ZHANG, ET AL. 2011. Measuring effective leaf area index, foliage profile, and stand height in New England forest stands using a full-waveform ground-based LiDAR. *Remote Sens. Environ.* 115:2954–2964.