Ecological Indicators 73 (2017) 156–165

Contents lists available at ScienceDirect

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind

# A comparison of cover calculation techniques for relating point-intercept vegetation sampling to remote sensing imagery☆

Jason W. Karl[a,*], Sarah E. McCord[b], Brian C. Hadley[c]

[a] USDA Agricultural Research Service, Jornada Experimental Range, Las Cruces, NM, USA
[b] Jornada Experimental Range, New Mexico State University, Las Cruces, NM, USA
[c] USDI Bureau of Land Management, National Operations Center, Denver, CO, USA

ARTICLE INFO

*Article history:*
Received 23 September 2015
Received in revised form 1 August 2016
Accepted 19 September 2016

*Keywords:*
Point-intercept sampling
Vegetation cover
Monitoring
Remote sensing
Beta regression
Method comparison

ABSTRACT

Accurate and timely spatial predictions of vegetation cover from remote imagery are an important data source for natural resource management. High-quality in situ data are needed to develop and validate these products. Point-intercept sampling techniques are a common method for obtaining quantitative information on vegetation cover that have been widely implemented in a number of local and national monitoring programs. The use of point-intercept data in remote sensing projects, however, is complicated due to differences in how vegetation cover indicators can be calculated. Decisions on whether to use plant intercepts from any canopy layer (i.e., any-hit cover) or only the first plant intercept at each point (i.e., top-hit cover) can result in discrepancies in cover estimates which are used to train remotely-sensed imagery. Our objective in this paper was to explore the theory of point-intercept sampling relative to training and testing remotely-sensed imagery, and to test the strength of relationships between top-hit and any-hit methods of calculating vegetation cover and high-resolution satellite imagery in two study areas managed by the Bureau of Land Management in northwestern Colorado and northeastern California. We modeled top-hit and any-hit percent cover for six vegetation indicators from 5m-resolution RapidEye imagery using beta regression. Model performance was judged using normalized root mean-squared error (RMSE) from a 5-fold cross validation. Any-hit cover estimates were significantly higher ($\alpha < 0.05$) than top-hit cover estimates for forbs and grasses in the White River study area, but only marginally higher in Northern California. Pseudo-$R^2$ values for beta regression models of vegetation cover from RapidEye image information varied from 0.1525 to 0.7732 in White River and 0.2455 to 0.6085 in Northern California, with little pattern to whether any-hit or top-hit indicators produced better model fit. However, normalized RMSE was lower for any-hit cover (indicating better model performance) or minimally higher than top-hit cover for all indicators in each study area. Our results do not support the idea that top-hit cover estimates from point-intercept sampling are the most appropriate for remote sensing applications in arid and semi-arid shrub-steppe environments. In fact, having two sets of different indicators calculated from the same data may cause additional confusion in a situation where there is already considerable debate on how vegetation cover should be measured and used. Ultimately, selection of indicators to use for developing remote sensing classification or predictive models should be based first on the meaning or interpretation of the indicator in the ecosystem of interest, and second on how well the indicator performs in modeling applications.



## 1. Introduction

Accurate and timely spatial predictions of vegetation cover are an important data source for natural resource management. Model-based approaches such as linear regression (e.g., McKenzie and Ryan, 1999), geostatistical techniques like regression kriging (e.g., Karl, 2010), regression model trees (e.g., Homer et al., 2012; Xian et al., 2015) or Bayesian regression trees (e.g., Harvey, 2011) are powerful techniques for predicting not only total vegetation cover, but also individual components of plant canopies. Developing such products at any scale, however, is an arduous and data hungry process, and relies heavily on in situ measurements to train, calibrate, and validate model predictions. The quality of spatial model predictions is largely dependent on availability of high-quality data that

☆ Mention of a proprietary product does not constitute a guarantee or warranty of the products by the U.S. Government or the authors and does not imply its approval to the exclusion of other products that may be suitable.

* Corresponding author at: USDA-ARS Jornada Experimental Range, P.O. Box 30003, MSC 3JER, New Mexico State University, Las Cruces, NM 88003-8003, USA.
*E-mail address:* jason.karl@ars.usda.gov (J.W. Karl).

http://dx.doi.org/10.1016/j.ecolind.2016.09.034
1470-160X/

are well-matched to remote imagery for model development and testing.

Point-intercept sampling techniques are a common method for obtaining high-quality quantitative information on vegetation cover (Bonham, 2013) that have been widely implemented in a number of local and national monitoring programs (e.g., Coulloudon et al., 1999; Herrick et al., 2009, 2010; Mackinnon et al., 2011). Point-intercept techniques estimate cover by the proportion of times vegetation is intercepted (i.e., touched) by either a pin or thin rod lowered, a laser beam shone, or cross hairs sighted through the plant canopy (all techniques collectively referred to as a "pin drop" for convenience). Point-intercept techniques can also be applied to very-high-resolution aerial imagery to produce estimates of vegetation cover (e.g., Booth and Cox, 2008). Studies have shown good correspondence between point-intercept vegetation cover data and various satellite and aerial image products (e.g., Laliberte et al., 2007; Karl and Maurer, 2010; Karl et al., 2014).

The use of point-intercept data in developing and testing models from remote-sensing imagery, however, is complicated due to differences in how vegetation cover indicators can be calculated. Decisions on whether to use plant intercepts from any canopy layer (i.e., any-hit cover) or only the first plant intercept at each point (i.e., top-hit cover) can result in discrepancies in cover estimates. Additionally, whether and how species are grouped into categories (e.g., life forms) for calculating cover can affect how well cover indicators correlate with image-based techniques (Booth and Cox, 2011).

There is little consistency in the literature on how vegetation cover indicators are calculated from point-intercept data. For example, Laliberte et al. (2007) used top-hit point-intercept indicators to predict vegetation cover from very-high resolution imagery. Seefeldt and Booth (2006) compared top-hit point-intercept estimates to cover estimates derived from interpretation of very-high resolution imagery. Serrano (2000) used top-hit point-intercept data to estimate cover of shrubs in a chaparral ecosystem. Alternatively, Karl (2010) and Karl et al. (2014) used any-hit point-intercept estimates to predict vegetation cover from satellite imagery. Knick and Rotenberry (2000) used any-hit point-intercept estimates to map bird habitats in southeastern Idaho. Many studies, however, do not provide enough detail on how point-intercept indicators are calculated or how different types of calculations are reconciled. For example Elmendorf et al. (2012) synthesized cover estimates from 61 studies that used a mix of top-hit and any-hit methods in arctic tundra.

Many applications of using in situ cover data to model plant cover from remote imagery appear to rest on the assumptions that only top canopy information is relevant to image analysis. McCoy (2005, p. 82) stated, "Unless a complete plant community composition analysis is needed, vegetation under the canopy may be ignored. Keep in mind that remote sensing field work is intended to evaluate what the satellite or aircraft sensor recorded. Cover density estimates are concerned with ground area covered from view above." However, the amount of radiation reflected from a land surface to a remote sensor is a complex function of the types and amounts of cover within and adjacent to an image pixel (Huang et al., 2002). Thus it is possible that lower canopy information from point-intercept vegetation sampling may be useful for remote sensing model development. This may be especially important in situations where the top canopy consists of many thin leaves and branches that may be frequently encountered in point-intercept sampling but not contribute much to overall pixel signature.

Thus the decision to take a top-canopy only view of point-intercept vegetation sampling data needs to appropriately consider the theory behind point-intercept sampling and to be empirically tested. Our objective in this paper was to explore the theory of point-intercept sampling relative to predicting vegetation cover indicators from remotely-sensed imagery, and to test the strength of relationships between high resolution satellite imagery and top-hit and any-hit methods of calculating vegetation cover in two study areas.

### *1.1. Theory of point-intercept sampling for vegetation cover*

For a detailed review of point-intercept sampling for vegetation cover, see Bonham (2013). Point sampling to estimate vegetation cover was first used by Levy (1927) and Levy and Madden (1933), and has become a commonly-used, objective methods for estimating vegetation cover (Bonham, 2013). The statistical properties of point-intercept sampling are described by Goodall (1953), Chen et al. (2008), and Bonham (2013).

Point-intercept sampling is based on the concept that if an infinite number of zero-dimensional points are placed in a two-dimensional area, the proportion of those points that intercept an object of interest (e.g., vegetation cover, a plant species, bare soil) equals the cover of that object within the defined area (Fig. 1, Bonham, 2013). Individual point intercepts are often treated as Bernoulli trials and cover of the object of interest ($\hat{p}$) is estimated as the sample mean of a binomial distribution:

$$\hat{p} = \frac{n_i}{n} \tag{1}$$

where $n_i$ is the number of intercepts where the object of interest was encountered, and $n$ is the total number of possible intercepts. Chen et al. (2008) showed that as $n$ approaches infinity, the proportion of $n_i/n$ asymptotically approaches the actual cover value of the object of interest.

Point-intercept sampling is often done at the intersections of lines within gridded quadrats or at regular intervals along a transect line. A pin or thin rod is lowered, a laser beam shone, or cross hairs sighted through the plant canopy (all techniques collectively referred to as a "pin drop" for convenience), and intercepting vegetation recorded, typically by species. Typically, all intercepting vegetation is recorded in the order in which it is encountered. Most commonly, each species is recorded only once per pin drop even if it is intercepted multiple times. In practice, several factors can influence the accuracy and precision of point-intercept estimates of vegetation cover. These include the total number of pin drops ($n$), configuration of transects or grid frames, angle of the pin drops, and diameter of the pin (see Elzinga et al., 1998; Bonham 2013).

Cover for a species (or life form) can be calculated either from the proportion of points where the species was the first intercepted vegetation for the pin drop (i.e., top-hit cover) or the proportion of points where the species was encountered at any position along the pin drop (i.e., any-hit cover). In the case of top-hit cover, cover percentages for all species (or life forms) will total 100%. For any-hit cover, the cover estimates reflect the actual cover of the species in the sampling area, but the percentages across species may total greater than 100%.

### *1.2. Study areas*

Comparisons of point-intercept cover calculations to satellite imagery were made on lands managed by the US Bureau of Land Management (BLM) in study areas in California and Colorado, USA (Fig. 2).

The White River study area in northwestern Colorado consisted of a 243,700-ha portion of the Piceance Basin managed by the BLM's White River Field Office (39.824° N, 108.297° W, Fig. 2). This area is characterized by deep valleys and high plateaus (Taylor, 1987). The BLM is the primary land steward in this area, managing approximately 70% of the study area. Vegetation in the Piceance Basin is mainly pinyon/juniper (*Pinus edulis* Engelm./*Juniperus osteosperma* Torr.) and sagebrush (*Artemisia* spp. L.) shrublands on the slopes and