A Comparison of Cover Calculation Techniques for Relating Point-Intercept Vegetation Sampling to Remote Sensing Imagery

Jason W. Karl¹, Sarah E. McCord², Brian C. Hadley³

¹ USDA Agricultural Research Service, Jornada Experimental Range, Las Cruces, New Mexico, USA

² Jornada Experimental Range, New Mexico State University, Las Cruces, New Mexico, USA

³ USDI Bureau of Land Management, National Operations Center, Denver, Colorado, USA

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:

Jason W. Karl

USDA-ARS Jornada Experimental Range

P.O. Box 30003, MSC 3JER

New Mexico State University

Las Cruces, NM 88003-8003

Phone: (575)-646-1301

Fax: (575)-646-5889

jkarl@nmsu.edu

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.

Keywords

Point-intercept sampling, vegetation cover, monitoring, remote sensing, beta regression, method comparison

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 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 et al. (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.

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 (Figure 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:

$$\hat{p} = \frac{n_i}{n} \quad (1)$$

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%.

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 (Figure 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, Figure 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 plateaus. Aspen (*Populus tremuloides* Michx.) woodlands are common in the southern portion of the study area. Drainage bottoms are typically cottonwood (*Populus* L.) and alder (*Alnus* Mill.) dominated riparian areas. Average annual precipitation in the White River study area from 1981 to 2010 was 443mm (data for the Meeker, CO weather station; National Oceanic and Atmospheric Administration, National Climate Data Center; <http://www.ncdc.noaa.gov/cdo-web/>; Accessed 22 September 2015).

The Northern California study area consisted of a 384,000-ha region of the Modoc Plateau region of northeast California (41.153° N, 120.145° W, Figure 2). The BLM is also the primary land steward, managing approximately 81% of the study area. This area is of basin and range geomorphology that results in a spectrum of ecosystems from desert playas through Great Basin sagebrush-steppe to high-elevation mountain big sagebrush (*Artemisia tridentata* Nutt. ssp. *vaseyana* (Rydb.) Beetle) communities. Average annual precipitation in the Northern California study area from 1981 to 2010 was 248mm (data for the Termo, CA weather station; National Oceanic and Atmospheric Administration, National Climate Data Center; <http://www.ncdc.noaa.gov/cdo-web/>; Accessed 22 September 2015).

Methods

Sample design and field data collection

Both of the study areas are part of long-term monitoring projects through the BLM's Assessment, Inventory and Monitoring (AIM) program (Mackinnon et al. 2011). Sample locations were randomly selected within the study areas in a manner to provide data to inform management objectives of the BLM's land use plans in each area. Because the objective of this

study was to examine the influence of different indicator calculations on the relationship of point-sampling data to satellite imagery, we used only that subset of sample locations from each study area that coincided with available imagery.

Sample selection in the White River study area was stratified by plant community potential as described in the area's soil surveys (U.S.D.A. Soil Conservation Service 1982, 1985, U.S.D.A. Natural Resources Conservation Service 2003). A total of 44 sample locations were selected across a range of different vegetation types. A more detailed description of sample site selection procedures for the White River area is in Karl et al. (2014).

Sample selection in the Northern California area was stratified using categories defined by an unsupervised classification of Landsat imagery because information on plant community potential was inconsistent (and in some cases unavailable) across the study area. The premise in using the Landsat classification for sample design was that it would help ensure a good distribution of sample locations across the different vegetation types in the study area. A total of 188 sample locations were selected for sampling in the Northern California AIM monitoring project in 2014. Of those, 101 fell within the bounds of the available satellite imagery for this study. For more information on the Northern California sample design, see BLM (2012).

At each sample location, point-intercept data for vegetation cover was collected according to the standard AIM line-point intercept protocol (Herrick et al. 2009, Mackinnon et al. 2011; for most recent version, see <http://www.landscapetoolbox.org/manuals/monitoring-manual/>, Accessed 18 September 2015). Field sampling in the White River study area occurred between 29 June and 13 September 2011. Sampling in the Northern California study area was between 13 May and 26 September 2014.

The point-intercept protocol consisted of three 50m transects radiating from the center of the sampling location (Herrick et al. 2009) at 120° intervals (Figure 2). The start of each transect was offset from the center by 5m to minimize the effects of trampling. At each meter along the transect, a thin (~1mm diameter) pin was lowered to the ground and all living and dead vegetation intercepting the pin was recorded in the order in which it was encountered. Each species was recorded only once per pin drop. Soil surface (e.g., mineral soil, rock, moss) or plant basal hits were also recorded for each pin drop. Herbaceous (<5mm diameter) and woody (>5mm diameter) plant litter was counted as a canopy layer. All plants were recorded to species except for sagebrush which were recorded to subspecies.

Cover indicator calculations

We followed Herrick et al. (2009) for the calculation of plant cover indicators from the point-intercept data. For the purposes of this study, we aggregated species-level information to major life forms, and calculated site-level indicator estimates for the following: annual and perennial forbs, annual and perennial grasses, shrubs, and sagebrush cover. Sagebrush cover was included in addition to the shrub life form indicator because of its importance to Greater Sagegrouse (*Centrocercus urophasianus*) habitat suitability (Connelly et al. 2000, Crawford et al. 2004).

At each sample location, transect data were pooled to calculate site-level cover estimates (i.e., the average across the three transects was assigned to the entire plot area). Any-hit cover was calculated for each life form as the total number of pin drops where that life form was encountered in any canopy layer divided by the total number of pin drops per sampling location

(n=150). Top-hit cover was calculated as the number of pin drops where the life form was the first plant encountered on the pin drop divided by the total number of pin drops.

Image acquisition and processing

RapidEye multispectral satellite imagery was acquired for the White River study area in 2010, and the Northern California study area in 2014. The imagery used for the White River study was for the year preceding field data collection due to high snow cover in northwestern Colorado during 2011 which obscured many of the field sites in the available RapidEye imagery for that year. The above-average precipitation the White River study area in 2011 could have resulted in less bare ground and greater cover of grasses and forbs that might affect correlations between the field data and imagery collected a year earlier. However, this would affect both methods of calculating the cover indicators (top-hit and any-hit) similarly, and as no significant changes in management or disturbance (e.g., fire) occurred at the sample locations between the image acquisition in 2010 and the field sampling in 2011, the use of the 2010 imagery was deemed acceptable. Additionally, Karl et al. (2014) found strong correlations between the 2011 field data and the 2010 RapidEye imagery despite the time difference.

The RapidEye satellite contains a 5-band multispectral sensor sampling the blue (440-510nm), green (520-590nm), red (630-685nm), red-edge (690-730nm), and near-infrared (760-850nm) wavelengths (<http://www.rapideye.com/products/ortho.htm>, date accessed 18 September 2015). The imagery was processed by RapidEye AG (Berlin, Germany) to their Level 3A which included radiometric and geometric corrections and resampling to a 5m ground-sampling distance (i.e., resolution) for 25 km² tiles. The White River study area consisted of five tiles acquired over two separate passes (June 15 and June 19, 2010). The Northern California study

area consisted of two tiles acquired over two separate passes (June 1 and June 4, 2014). Image tiles for each study area were merged together using histogram matching on overlapping areas.

The RapidEye satellites can acquire imagery at off-nadir angles. For White River, images were collected at 6.78° and 6.42° off-nadir for the June 15 and June 19, 2010 acquisitions, respectively. For Northern California, images were collected at 9.85° and 3.07° off-nadir for the June 1 and June 4, 2014 acquisitions, respectively. Off-nadir view angles can affect the radiance values recorded at the sensor because reflected radiation must travel through more plant canopy. Myneni and Williams (1994) found that vegetation indices increased with off-nadir image angle, however, the impact at low off-nadir angles ($< 10^\circ$) was minor (see Myneni and Williams, 1994, fig. 3). Additionally, effects of off-nadir view angle would be consistent across the image, and should not significantly affect statistical relationships between field measurements and imagery values. For these reasons, we concluded that any effect on image values induced by off-nadir view angle was negligible.

For each field sample location, we calculated the average and standard deviation of the pixel values for each RapidEye band within the same plot radius defined by the field transects. Because the locations of the transects within the plot was not precisely known, and because the field transects were intended only to spread field observations across the plot and not as formal plot sub-samples (Mackinnon et al. 2011), it was not possible to assign specific RapidEye pixels to individual transects or portions of transects. Averaging RapidEye band values within the field plot area provided a set of imagery measures commensurate in scale with the field observations.

Statistical analyses

Statistical analysis followed two steps: First, we evaluated the correspondence between the any-hit and top-hit cover calculations for each indicator. We then used a beta-distribution regression to quantify the relationship between any-hit and top-hit cover indicators and the RapidEye imagery. All statistical analyses were performed in R version 3.1 (<http://www.r-project.org>, accessed 1 September, 2015) using the *betareg* package developed by Cribari-Neto and Zeileis (2010) for the beta-distribution regression.

Discrepancies in how any-hit and top-hit cover indicators correlate with remote imagery will depend in part on how strongly related the two indicator calculations are to each other. Accordingly, we first plotted the top-hit and any-hit cover estimates against each other and calculated the mean and sample standard deviation for each indicator. We tested for differences between the any-hit and top-hit cover calculations for each indicator using a t-test and used an F test to determine differences in variance between any-hit and top-hit cover indicator estimates.

Vegetation cover data (expressed as percentages or proportions) often do not conform to the assumptions of standard linear regressions because cover values are constrained by the interval $[0,1]$, are typically not normally-distributed, and can exhibit variances that change over the range of values (i.e., variance approaches zero at the limits of the range) (Espinheira et al. 2008, Damgaard 2009). An alternative for analyzing and modeling cover data is to use a beta-distribution regression (Brehm and Gates 1993, Bonham 2013). Chen et al. (2008) showed that plant cover data follow a beta distribution, and Damgaard (2009) proposed that statistics based on a beta distribution were preferable to those using a normal distribution for continuous plant cover data. For a thorough discussion and formulation of the beta distribution and beta regression, see Ferrari and Cribari-Neto (2004) and Chen et al. (2008).

Beta regression is useful for analyzing plant cover data for three reasons. First, the beta distribution is bounded by 0 and 1 and assumes only that the variable being analyzed is continuous and bounded between two known endpoints (Smithson and Verkuilen 2006). Both criteria are true for plant cover data. Second, the beta distribution can take on many different shapes depending on the value of its two parameters: the mean (μ) and precision (ϕ). Additionally, variance of a beta distribution is a function of both μ and ϕ ,

$$\sigma^2 = \frac{\mu(1-\mu)}{(1+\phi)} \quad (2)$$

which allows the variance to be modeled independently from the mean (Smithson and Verkuilen 2006, Ospina and Ferrari 2011) and accurately describes observed variances of plant cover data (e.g., variance is smaller at the tails of the distribution). Third, because the beta distribution can take on many different shapes, cover data need not be transformed for analysis even if the data are highly skewed (Espinheira et al. 2008).

We used beta regression to model the relationship between the two cover indicator calculations and the average and standard deviation RapidEye band values at each sample location. The assumption with this approach is that the most appropriate method of calculating an indicator will result in a model that explains more of the variability in the indicator values. Beta regression produces a pseudo- R^2 value as a measure of explanatory power of the model (Ferrari and Cribari-Neto 2004). While pseudo- R^2 values are an indication of model fit, they can be subject to overfitting and are subsequently not a good measure of overall model performance. To measure model performance, we performed 1,000 iterations of a 5-fold cross-validation (i.e., leave five randomly selected points out, construct the model and use it to predict the values of the withheld data) to produce a root mean-squared error (RMSE) for the cover indicator predictions. Because RMSE values are only directly comparable if the distributions being

compared are similar, we normalized RMSE by dividing our calculated RMSE values for each model by the standard deviation of the indicator being predicted. Lower normalized RMSE values indicate better model performance. Overall performance of any-hit and top-hit calculations for each cover indicator was based jointly on the model fit and the normalized RMSE of the predictions. The beta-regression models for any-hit and top-hit calculations for each indicator were applied back to the original RapidEye imagery to produce maps of cover values across the study areas. These maps were used to explore how the different indicator calculations affect the spatial distribution of cover predictions.

Results

Total foliar cover across sample sites in the White River study area averaged 48.8% (sd = 16.9%), whereas average total foliar cover for the Northern California sites was 69.4% (sd = 21.7%) (Figure 3).

Any-hit cover estimates were significantly higher (at $\alpha = 0.05$ level) than top-hit cover estimates for perennial and annual forbs and grasses in the White River study area (Table 1, Figure 4). Variances of the any-hit cover estimates also were significantly greater than variances of top-hit cover for annual and perennial forbs and grasses in this study area ($p < 0.01$ for all four indicators). For the shrub and sagebrush indicators, top-hit cover estimates were only slightly lower than any-hit estimates, and variances were slightly lower but not significantly different ($p=0.56$ and $p=0.4515$ for shrub and sagebrush indicators, respectively) between the two methods.

In the Northern California study area, differences between any-hit and top-hit cover estimates by indicator were generally minimal (Table 1, Figure 5), with only perennial forbs

showing a modestly significant difference ($p=0.099$). Variances of the any-hit and top-hit cover estimates were significantly different only for annual ($p=0.085$) and perennial forbs ($p=0.007$).

For the White River study area, beta regression model fit varied among indicator and method of calculating cover from $\text{pseudo-R}^2 = 0.1525$ for top-hit cover of perennial forbs to $\text{pseudo-R}^2 = 0.7732$ for any-hit cover of shrubs (Table 2). Any-hit cover yielded higher pseudo-R^2 for perennial forbs, perennial grasses, and shrubs. Top-hit cover gave slightly higher pseudo-R^2 values for annual forbs and annual grasses. Any-hit and top-hit Pseudo-R^2 values were similar for the sagebrush indicator. Any-hit cover models in the White River study area yielded lower normalized RMSE for all indicators except annual forbs and sagebrush (though the difference was slight for these indicators). When mapped to the White River study area, the any-hit cover models predicted a greater area as having higher cover values than the top-hit cover models for all indicators (Figure 6).

In the Northern California study area, beta regression model fit also varied among indicator and method of calculating cover from $\text{pseudo-R}^2 = 0.2455$ for top-hit perennial grass cover to $\text{pseudo-R}^2 = 0.6085$ for top-hit annual grass cover. Overall, however, differences in model fit and overall model performance for Northern California were smaller than those found in White River (Table 2). Any-hit cover produced higher pseudo-R^2 values for perennial forbs and grasses and for sagebrush; whereas top-hit cover produced higher pseudo-R^2 values for annual forbs and grasses and all shrubs. Any-hit cover models in Northern California yielded lower normalized RMSE than top-hit cover models for the annual forb, perennial forb, and perennial grass indicators. Top-hit cover models gave lower normalized RMSE for the annual grass, shrub, and sagebrush indicators. However, for all indicators except perennial forbs and perennial grass, the difference in normalized RMSE was relatively small. Mapped model

predictions for the Northern California study area also showed little difference between the any-hit and top-hit cover models (Figure 7).

Discussion

In both of our study areas, any-hit cover estimates performed as well as or better than top-hit estimates for predicting cover from RapidEye imagery for all indicators. This result contradicts the conventional idea that only top-canopy information is relevant for remote sensing applications (e.g., McCoy 2005). In part, this may be due to differences in scale and definition of cover between point-intercept sampling and remote imagery. Point-intercept techniques estimate cover over a defined area by the proportion of (theoretically zero-dimensional) points where something is encountered. While the proportion of that cover in the top canopy versus lower canopy is a legitimate indicator to calculate from these data, point-intercept does so using a strictly foliar definition of canopy, where canopy consists of any plant material, no matter how small, that covers the ground surface when viewed from an aerial perspective (Bedell 1998). This definition of canopy may not always correspond with the objects inside an image pixel that reflect radiation and contribute to the values recorded at the sensor. For example, in point-intercept sampling, it is not uncommon to intercept a thin flowering stalk of a grass before intercepting more substantial vegetation below it. In this case the grass would be counted as the top canopy when in fact the vegetation below it most likely contributes more to the value recorded by the remote sensor.

Field data from the White River study area showed a greater variation in differences between any-hit and top-hit cover calculations than did the data from Northern California (Figures 4 and 5). The mechanical explanation for this is that field crews collecting point-

intercept in White River recorded more lower-canopy intercepts than the crews collecting data in Northern California (see Figure 1). There are several reasons why this could occur. First, differences in overall productivity between the two study areas (evidenced by distributions of total foliar cover of sampling sites, Figure 3) may lead to more vegetation being encountered at each pin drop in White River. Second, differences in management and disturbance history between the study areas may have affected plant community composition and structure at the sample locations. In other words, the Northern California study area had, on average, a simpler vegetation structure (i.e., fewer canopy layers) than White River. A third reason could be differences in the ways that the two field crews were trained or in which they implemented the line-point intercept method that introduced bias into the data. For example, crews in White River may have held the intercept pin at an angle which would result in more plant intercepts (see Bonham 2013). This third reason is likely not a significant cause of the differences as field data from the study areas in other years (2012-2014 for White River; 2013 for Northern California) showed similar patterns.

Regardless of the cause, the implication of the difference in magnitude between cover estimates among study areas is that there is a greater potential for any-hit and top-hit calculations to yield different results for remote sensing models and classifications. This was borne out in our results where the normalized RMSE values and mapped cover predictions for Northern California were similar between methods but different in many cases in White River.

Any-hit and top-hit estimates of vegetation cover from point-intercept sampling are not the same indicators – they measure different aspects of the plant community. While the two measures may be very similar for plants typically encountered in the upper canopy or for total-cover indicators, they can be very different for other aspects. Regardless of which method is

better correlated to remotely-sensed imagery, the two metrics may not be interchangeable, especially for plants frequently occurring in the understory. Selection of which method to use should be based upon the ecological interpretation of the indicators. For example, Sage-grouse (*Centrocercus urophasianus*) habitat suitability is determined, in part, by presence and abundance of certain forbs and grasses (Connelly et al. 2000, Crawford et al. 2004, Stiver et al. 2015). In this case, predictions of top-hit forb cover (or abundance) may not equate well to Sage-grouse habitat suitability.

This study presents a comparison of different methods for calculating vegetation cover from point-intercept sampling data for a limited set of indicators, a single remote-sensing source, and in only two study systems. Even though our results were consistent among our study areas, further investigation is warranted into the utility of different techniques for calculating cover indicators for different applications. For example, in denser canopy environments (e.g., forests) would any-hit cover indicators still perform equal to or better than top-hit cover indicators? We encourage researchers to compare different methods of calculating indicators as a matter of practice in indicator selection for developing remote-sensing products.

Conclusion

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 (e.g., Whitman and Siggeirsson 1954, Johnston 1957, Hatton et al. 1986, Floyd and Anderson 1987, Seefeldt and Booth 2006, Bonham

2013, Karl et al. 2014, Thacker et al. 2015). 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 for a given ecosystem, and second on how well the indicator performs in modeling applications.

Acknowledgements

We would like to express appreciation to the BLM Assessment, Inventory, and Monitoring field crews who collected the field data in the White River and Northern California study areas. We also thank S.F. Perry for her assistance in compiling the data and calculating the indicators used in this study. Rapid Eye imagery was provided by the BLM's National Operations Center. Support for this work was provided by the BLM Assessment, Inventory, and Monitoring program, and appropriated funds to the USDA-ARS Range Research Unit, Jornada Experimental Range.

References

- Bedell, T. E. 1998. Glossary of terms used in range management. 4th Edition. Society for Range Management. Direct Press, Denver, Colorado.
- Bonham, C. D. 2013. Measurements for terrestrial vegetation. Second Edition. Wiley-Blackwell, Chichester, West Sussex ; Hoboken, NJ.
- Booth, D. T., and S. E. Cox. 2011. Art to Science: Tools for Greater Objectivity in Resource Monitoring. *Rangelands* 33:27–34.
- Booth, T. D., and S. E. Cox. 2008. Image-based monitoring to measure ecological change in rangeland. *Frontiers of Ecology and the Environment* 6:185–190.

- Brehm, J., and S. Gates. 1993. Donut shops and speed traps: Evaluating models of supervision on police behavior. *American Journal of Political Science* 37:555–581.
- Bureau of Land Management. 1996. Sampling vegetation attributes: interagency technical reference. BLM National Applied Resource Sciences Center.
- Bureau of Land Management. 2012. Rush fire emergency stabilization and rehabilitation environmental assessment. Page 128. Bureau of Land Management, Susanville, CA.
- Chen, J., M. Shiyomi, C. Bonham, T. Yasuda, Y. Hori, and Y. Yamamura. 2008. Plant cover estimation based on the beta distribution in grassland vegetation. *Ecological Research* 23:813–819.
- Connelly, J. W., M. A. Schroeder, A. R. Sands, and C. E. Braun. 2000. Guidelines to manage sage grouse populations and their habitats. *Wildlife Society Bulletin* 28:967–985.
- Crawford, J. A., R. A. Olson, N. E. West, J. C. Mosley, M. A. Schroeder, T. D. Whitson, R. F. Miller, M. A. Gregg, and C. S. Boyd. 2004. Ecology and management of sage-grouse and sage-grouse habitat. *Journal of Range Management* 57:2–19.
- Cribari-Neto, F., and A. Zeileis. 2010. Beta regression in R. *Journal of Statistical Software* 34:1–24.
- Damgaard, C. 2009. On the distribution of plant abundance data. *Ecological Informatics* 4:76–82.
- Elmendorf, S. C., G. H. R. Henry, R. D. Hollister, R. G. Björk, A. D. Bjorkman, T. V. Callaghan, L. S. Collier, E. J. Cooper, J. H. C. Cornelissen, T. A. Day, A. M. Fosaa, W. A. Gould, J. Grétarsdóttir, J. Harte, L. Hermanutz, D. S. Hik, A. Hofgaard, F. Jarrad, I. S. Jónsdóttir, F. Keuper, K. Klanderud, J. A. Klein, S. Koh, G. Kudo, S. I. Lang, V. Loewen, J. L. May, J. Mercado, A. Michelsen, U. Molau, I. H. Myers-Smith, S. F. Oberbauer, S. Pieper, E.

- Post, C. Rixen, C. H. Robinson, N. M. Schmidt, G. R. Shaver, A. Stenström, A. Tolvanen, Ø. Totland, T. Troxler, C.-H. Wahren, P. J. Webber, J. M. Welker, and P. A. Wookey. 2012. Global assessment of experimental climate warming on tundra vegetation: heterogeneity over space and time: Warming effects on tundra vegetation. *Ecology Letters* 15:164–175.
- Elzinga, C. L., D. W. Salzer, and J. W. Willoughby. 1998. Measuring and monitoring plant populations. U.S. Department of the Interior, Bureau of Land Management. National Applied Resource Sciences Center, Denver, Colorado.
- Espinheira, P. L., S. L. P. Ferrari, and F. Cribari-Neto. 2008. On beta regression residuals. *Journal of Applied Statistics* 35:407–419.
- Ferrari, S., and F. Cribari-Neto. 2004. Beta regression for modelling rates and proportions. *Journal of Applied Statistics* 31:799–815.
- Floyd, D. A., and J. E. Anderson. 1987. A comparison of three methods for estimating plant cover. *Journal of Ecology* 71:221–228.
- Goodall, D. 1953. Point-quadrat methods for the analysis of vegetation. *Australian Journal of Botany* 1.
- Harvey, J. W. 2011. Using Bayesian Regression Tree Models and remotely sensed data to characterize recent environmental change in Alaska, USA. Rutgers, The State University of New Jersey, New Brunswick, New Jersey.
- Hatton, T. J., N. E. West, and P. S. Johnson. 1986. Relationships of the error associated with ocular estimation and actual total cover. *Journal of Range Management* 39:91–92.

- Herrick, J. E., V. Lessard, K. E. Spaeth, P. Shaver, R. S. Dayton, D. A. Pyke, L. Jolley, and J. J. Goebel. 2010. National ecosystem assessments supported by scientific and local knowledge. *Frontiers of Ecology and the Environment* 8:403–408.
- Herrick, J. E., J. W. Van Zee, K. M. Havstad, L. M. Burkett, and W. G. Whitford. 2009. Monitoring manual for grassland, shrubland, and savanna ecosystems. USDA-ARS Jornada Experimental Range, Las Cruces, New Mexico.
- Homer, C. G., C. L. Aldridge, D. K. Meyer, and S. J. Schell. 2012. Multi-scale remote sensing sagebrush characterization with regression trees over Wyoming, USA: Laying a foundation for monitoring. *International Journal of Applied Earth Observation and Geoinformation* 14:233–244.
- Huang, C., J. R. G. Townshend, S. Liang, S. N. V. Kalluri, and R. S. DeFries. 2002. Impact of sensor's point spread function on land cover characterization: assessment and deconvolution. *Remote Sensing of Environment* 80:203–212.
- Johnston, A. 1957. A COMPARISON OF THE LINE INTERCEPTION, VERTICAL POINT QUADRAT, AND LOOP METHODS AS USED IN MEASURING BASAL AREA OF GRASSLAND VEGETATION. *Canadian Journal of Plant Science* 37:34–42.
- Karl, J. W. 2010. Spatial predictions of cover attributes of rangeland ecosystems using regression kriging and remote sensing. *Rangeland Ecology and Management* 63:335–349.
- Karl, J. W., and B. A. Maurer. 2010. Multivariate correlations between imagery and field measurements across scales: comparing pixel aggregation and image segmentation. *Landscape Ecology* 24:591–605.

- Karl, J. W., J. Taylor, and M. Bobo. 2014. A double-sampling approach to deriving training and validation data for remotely-sensed vegetation products. *International Journal of Remote Sensing* 35:1936–1955.
- Knick, S. T., and J. T. Rotenberry. 2000. Ghosts of habitats past: contribution of landscape change to current habitats used by shrubland birds. *Ecology* 81:220–227.
- Laliberte, A. S., A. Rango, J. E. Herrick, E. L. Fredrickson, and L. Burkett. 2007. An object-based image analysis approach for determining fractional cover of senescent and green vegetation with digital plot photography. *Journal of Arid Environments* 69:1–14.
- Levy, E. 1927. The grasslands of New Zealand. *New Zealand Journal of Agriculture* 34:145–164.
- Levy, E., and E. Madden. 1933. The point method for pasture analysis. *New Zealand Journal of Agriculture* 46:267–279.
- Mackinnon, W. C., J. W. Karl, G. R. Toevs, J. J. Taylor, M. S. Karl, C. S. Spurrier, and J. E. Herrick. 2011. BLM core terrestrial indicators and methods. US Department of the Interior, Bureau of Land Management, National Operations Center, Denver, CO.
- McCoy, R. M. 2005. *Field methods in remote sensing*. The Guilford Press, New York, NY.
- McKenzie, N., and P. Ryan. 1999. Spatial prediction of soil properties using environmental correlation. *Geoderma* 89:67–94.
- Myneni, R. B., and D. L. Williams. 1994. On the relationship between FAPAR and NDVI. *Remote Sensing of Environment* 49:200–211.
- Ospina, R., and S. L. P. Ferrari. 2011. A general class of zero-or-one inflated beta regression models. *Computational Statistics & Data Analysis* 56:1609–1623.

- Seefeldt, S. S., and T. D. Booth. 2006. Measuring plant cover in sagebrush steppe rangelands: a comparison of methods. *Environmental Management* 37:703–711.
- Serrano, L. 2000. Deriving Water Content of Chaparral Vegetation from AVIRIS Data. *Remote Sensing of Environment* 74:570–581.
- Smithson, M., and J. Verkuilen. 2006. A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. *Psychological Methods* 11:54–71.
- Stiver, S. J., E. T. Rinkes, D. E. Naugle, P. D. Makela, D. A. Nance, and J. W. Karl. 2015. Sage-grouse habitat assessment framework: a multiscale assessment tool. Page 114. Technical Reference, Bureau of Land Management and Western Association of Fish and Wildlife Agencies, Denver, CO.
- Taylor, J. O. 1987. Oil shale, water resources, and valuable minerals of the Piceance Basin, Colorado; the challenge and choices of development. Page 143.
- Thacker, E., T. Messmer, and B. Burritt. 2015. Sage-Grouse Habitat Monitoring: Daubenmire versus Line-Point Intercept. *Rangelands* 37:7–13.
- U.S.D.A. Natural Resources Conservation Service. 2003. Soil survey of Douglas-Plateau area, Colorado, parts of Garfield and Mesa counties. Page 356. Washington, D.C.
- U.S.D.A. Soil Conservation Service. 1982. Soil survey of Rio Blanco county area, Colorado. Page 164. Washington, D.C.
- U.S.D.A. Soil Conservation Service. 1985. Soil survey of Rifle area, Colorado. Page 161. Washington, D.C.
- Whitman, W. C., and E. I. Siggeirsson. 1954. Comparison of Line Interception and Point Contact Methods in the Analysis of Mixed Grass Range Vegetation. *Ecology* 35:431.

Xian, G., C. Homer, M. Rigge, H. Shi, and D. Meyer. 2015. Characterization of shrubland ecosystem components as continuous fields in the northwest United States. *Remote Sensing of Environment* 168:286–300.

Table 1. Summary statistics for any-hit and top-hit cover indicators in the White River and Northern California study areas. Differences in means and variances between any-hit and top-hit indicators were tested using a t-test on sample means and a F-test on sample variances, respectively.

<u>White River</u>						
	Any-hit Cover		Top-hit Cover		t-test	F-Test
	Mean	Std. Dev.	Mean	Std. Dev.		
Annual Forb	1.67%	2.06%	0.56%	0.95%	3.183***	4.679***
Perennial Forb	16.42%	19.68%	5.60%	4.34%	3.561***	20.528***
Annual Grass	1.39%	2.28%	0.55%	0.94%	2.280**	5.857***
Perennial Grass	34.78%	22.36%	14.25%	9.92%	5.567***	5.077***
Shrubs	39.78%	24.69%	32.53%	22.58%	1.436	1.196
Sagebrush	13.60%	13.95%	11.84%	12.43%	0.624	1.260
<u>Northern California</u>						
	Any-hit Cover		Top-hit Cover		t-test	F-Test
	Mean	Std. Dev.	Mean	Std. Dev.		
Annual Forb	4.13%	5.53%	3.34%	4.66%	1.093	1.413*
Perennial Forb	3.58%	4.93%	2.56%	3.75%	1.660*	1.162**
Annual Grass	26.22%	18.03%	23.00%	16.46%	1.324	1.201
Perennial Grass	13.04%	12.67%	11.49%	12.02%	0.897	1.120
Shrubs	8.16%	9.43%	7.60%	9.15%	0.429	1.062
Sagebrush	5.13%	7.35%	4.84%	6.96%	0.282	1.114

* denotes significant at $\alpha = 0.1$; ** denotes significant at $\alpha = 0.05$; *** denotes significant at $\alpha < 0.01$.

Table 2. Beta-model regression results for White River and Northern California study areas.

Pseudo R^2 is a statement of model fit for predicting cover indicators from RapidEye satellite imagery. RMSE and nRMSE are root mean-squared error and normalized root mean-squared error from a model cross validation.

<u>White River</u>						
	Any-hit Cover			Top-hit Cover		
	Pseudo R^2	RMSE	nRMSE	Pseudo R^2	RMSE	nRMSE
Annual Forb	0.1772	0.0216	1.0469	0.1942	0.0096	1.0044
Perennial Forb	0.3544	0.1896	0.9638	0.1525	0.0481	1.1076
Annual Grass	0.3122	0.0215	1.0582	0.3225	0.0111	1.1770
Perennial Grass	0.6490	0.1363	0.6097	0.2763	0.0946	0.9536
Shrubs	0.7732	0.1445	0.5852	0.6700	0.1606	0.7112
Sagebrush	0.4833	0.1218	0.8728	0.4785	0.1072	0.8629
<u>Northern California</u>						
	Any-hit Cover			Top-hit Cover		
	Pseudo R^2	RMSE	nRMSE	Pseudo R^2	RMSE	nRMSE
Annual Forb	0.2888	0.0501	0.9270	0.3094	0.0449	0.9889
Perennial Forb	0.3496	0.0353	0.7129	0.3243	0.0316	0.8353
Annual Grass	0.5499	0.1138	0.6324	0.6085	0.0997	0.6081
Perennial Grass	0.2739	0.0990	0.8010	0.2455	0.1051	0.8973
Shrubs	0.3509	0.0827	0.8906	0.3995	0.0773	0.8568
Sagebrush	0.3504	0.0815	1.0782	0.3434	0.0752	1.0482

Figure 1. Line-point intercept is an implementation of general point-intercept sampling for estimating vegetation cover at a site. A slender pin or metal rod (~1mm diameter) is lowered through the plant canopy (A) and vegetation touching (i.e., intercepting) the pin is recorded in the order in which it was intercepted (B and C). Cover can be calculated as either the proportion of positions along the transect where a particular species or life form was encountered in any canopy layer (i.e., any-hit cover) or the proportion of positions along the transect where the species was encountered as the top canopy layer (i.e., top-hit cover). Figure adapted from Herrick et al. (2009).

Figure 2. Location of the White River and Northern California study areas and configuration of the vegetation sampling plots. Study area boundaries are coincident with RapidEye imagery used for the analysis.

Figure 3. Histograms of total foliar cover for sample sites in the White River (n=44) and Northern California (n=101) study areas.

Figure 4. Correlation between any-hit and top-hit cover calculations for the six life form indicators in the White River study area.

Figure 5. Correlation between any-hit and top-hit cover calculations for the six life form indicators in the Northern California study area.

Figure 6. Examples of differences between any-hit and top-hit cover model predictions for shrub and perennial grass cover in the White River study area.

Figure 7. Examples of differences between any-hit and top-hit cover model predictions for shrub and perennial grass cover in the White River study area.