Remote Sensing of Environment xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Combining airborne hyperspectral and LiDAR data across local sites for upscaling shrubland structural information: Lessons for HyspIRI

Jessica J. Mitchell [a,*], Rupesh Shrestha [b], Lucas P. Spaete [b], Nancy F. Glenn [b]

[a] Department of Geography and Planning, Appalachian State University, P.O. Box 32066, Boone, NC 28608, USA
[b] Boise Center Aerospace Laboratory, Department of Geosciences, Boise State University, 1910 University Drive, Boise, ID 83725-1535, USA

ARTICLE INFO

Article history:
Received 1 June 2014
Received in revised form 2 April 2015
Accepted 5 April 2015
Available online xxxx

Keywords:
LiDAR
Hyperspectral
HsypIRI
Sagebrush
Vegetation structure

ABSTRACT

Fine-scale variation of vegetation structure in dryland systems, such as the Great Basin in the western US, is critical to understanding ecosystem responses to changing land-use conditions. High resolution airborne hyperspectral (HyMap) and LiDAR datasets acquired across independent collection sites can reduce uncertainty in predictive ecosystem modeling and provide a basis for regional upscaling to satellite observations of structural metrics such as cover and height. In the first part of our study, we combined ground reference and airborne data collected at three sagebrush-steppe locations and used the statistical data mining tool random forests to identify remote sensing variables most relevant to estimating shrub cover. In the second part of our study, we hypothesized that vegetation indices derived from hyperspectral satellite observations would not only reliably predict shrub cover but also be relatable to shrub height; thereby augmenting the collection of vertical structure estimates from future satellite platforms such as ICESAT-2. To test this hypothesis, we simulated HyspIRI observations to derive variables to relate to LiDAR-based estimates of shrub cover and height. We generated the same hyperspectral variables as in the first part of this study but at coarser resolution (60 m) and we again used random forests to model shrub cover and height and identify predictors of greatest importance. Overall, combining LiDAR and HyMap datasets at the airborne scale improved shrub cover model results ($r^2 = 0.58$) compared to LiDAR alone ($r^2 = 0.49$). Primary shrub cover variables of importance were H_{IQR} (the interquartile range of height of all LiDAR vegetation returns), H_{MAD} (median absolute deviation from median height of all LiDAR vegetation returns), a narrowband index sensitive to anthocyanins, the ratio of LiDAR vegetation returns to total returns, and a red to green ratio. In addition, HyspIRI-simulated narrowband vegetation indices were relatable to LiDAR-derived shrub cover and height variables (r^2 ranging from 0.63 to 0.71) with relatively low root mean square error.

1. Introduction

Sagebrush (*Artemisia* spp.) communities once covered approximately 63 million ha of rangeland in the western United States and Canada and represent the largest and one of the most threatened ecosystems in the temperate semi-desert ecoregion of North America (Anderson & Inouye, 2001; Homer, Aldridge, Meyer, & Schell, 2012). Sagebrush habitat provides food or cover for over 350 wildlife species including sage grouse (Knick & Connelly, 2011; Suring, Rowland, & Wisdom, 2005; Tilley, Ogle, John, & Benson, 2006). Like most vegetation, sagebrush cover and height characteristics vary across the landscape. Accurately mapping this variation is important for sage grouse habitat selection, which depends on percent canopy cover, visual cover and height; and for habitat modeling (e.g. Crawford et al., 2004; Krogh, Zeisset, Jackson, & Whitford, 2002). Cover and height are also relevant to estimating fuel loads (e.g. Castedo-Dorado, Gómez-Vázquez, Fernandes, & Crecente-Campo, 2012; Keane, Rollins, McNicoll, & Parsons, 2002) and aboveground biomass (Mathieu et al., 2013), which are indicators of forage potential, species dominance and hydrologic function in semi-arid systems. When coupled with canopy shape, sagebrush cover and height provide information about the spatial pattern of vegetation roughness, which directly affects aeolian sediment transport (Mueller, Wainwright, & Parsons, 2007; Okin, 2008) and may be relatable to aerodynamic roughness, a key parameter in energy balance models and evapotranspiration (Lee, Timmermans, Su, & Mancini, 2012) and shrub patch dynamics (Schlesinger et al., 1990). Fine-scale characterization of the variability in sagebrush height and cover is important to initialize terrestrial ecosystem models (e.g. Medvigy, Wofsy, Munger, Hollinger, & Moorcroft, 2009) to understand structural dynamics and provide regional estimates of carbon stock and fluxes under future climate change scenarios.

Several studies have demonstrated the use of multispectral imagery (1 m to 56 m pixels) for monitoring categorical and continuous shrub cover change in sagebrush ecosystems (e.g., Ramsey, Wright, & McGinty, 2004; Sivanpallai, Prager, & Storet, 2009; Stow, Hamada,

* Corresponding author. Tel.: +1 828 262 7054.
E-mail address: mitchelljj@appstate.edu (J.J. Mitchell).

http://dx.doi.org/10.1016/j.rse.2015.04.015
0034-4257/

Coulter, & Anguelova, 2008). However, multispectral and hyperspectral studies designed to estimate vegetation cover in sagebrush are limited by multiple scattering, bright soil reflectance, penetrable canopies and spectrally indiscriminate targets (e.g., Laliberte, Fredrickson, & Rango, 2007; Mitchell & Glenn, 2009; Okin, Roberts, Murray, & Okin, 2001; Smith, Ustin, Adams, & Gillespie, 1990). Small-footprint, discrete return light detection and ranging (LiDAR), or airborne laser scanning, is not limited by many of these spectral challenges; however, separating LiDAR returns in low-height, open canopy rangeland vegetation is difficult because the vegetation canopy returns are often close to ground returns. Recent studies confirm the appropriateness of LiDAR for structural and biomass applications (Latifi, Fassnacht, & Koch, 2012; Swatantran, Dubayah, Roberts, Hofton, & Blair, 2011; Zolkos, Goetz, & Dubayah, 2013), with hyperspectral data providing important canopy stress information (Swatantran et al., 2011) and minor improvements to the LiDAR models (e.g., Anderson et al., 2008; Latifi et al., 2012; Mundt, Streutker, & Glenn, 2006). While combining LiDAR-derived estimates of vegetation structure with hyperspectral information tends to result in slightly improved accuracy, new methods are needed to optimize these datasets; understand the relative tradeoffs and redundancies between the two sensors; identify uncertainties associated with upscaling; and develop composite products that can be iteratively assessed and refined in terms of prediction accuracy (Esteban, Starr, Willetts, Hannah, & Bryanston-Cross, 2005).

Furthermore, an improved understanding of the contribution of hyperspectral data in estimating vegetation structure will improve future applications of hyperspectral infrared imager (HyspIRI) data, along with synergistic use of HypsIRI with other remote sensing data, such airborne hyperspectral and LiDAR, and ICESat-2's advanced topographic laser altimeter (ATLAS). HyspIRI is a future National Research Council (NRC) decadal survey mission from National Aeronautics and Space Administration (NASA) that is expected to be launched in the next decade (NRC, 2007). One of the instruments onboard HyspIRI is an imaging spectrometer yielding 60 m spatial resolution data in 10 nm contiguous bands ranging from 380 nm to 2500 nm at an equatorial 19 day repeat cycle (NASA, 2014). The spectral range and bandwidth are similar to that of the Hyperion sensor on NASA's EO-1 satellite. Hyperion can collect transect samples in narrow swaths at 30 m spatial resolution but suffers from cross-track calibration issues and is limited by low signal-to-noise (Pearlman et al., 2003). In contrast, HyspIRI is a global imager and the mission is primarily expected to contribute to our understanding of carbon and ecosystem processes by enabling global vegetation mapping at finer taxonomic levels and rapid detection of plant stresses. Recently, using simulated data, various studies have demonstrated the potential of HyspIRI in different applications such as vegetation mapping (Olsson & Morisette, 2014), estimation of fraction of photosynthetically active radiation and leaf water content (Zhang, Middleton, Gao, & Cheng, 2012), and in other geoscience (Abrams, Pieri, Realmuto, & Wright, 2013; Kruse et al., 2011) and urban applications (Roberts, Quattrochi, Hulley, Hook, & Green, 2012). Similar studies on more complex shrubland ecosystems can provide insights into the potential of HyspIRI in estimating vegetation structural parameters such as cover and biomass. Airborne hyperspectral data obtained from NASA's AVIRIS sensor (limited availability due to commissioning requirement) or commercial instruments such as HyMap (HyVista Co., Sydney, Australia) contain similar spectral coverage and can be relevant proxies to generate such simulations.

This study analyzes and integrates HyMap and LiDAR data using a random forest approach (Breiman, 2001), which can be used to select (indirectly) important predictor variables and has been demonstrated to predict forest canopy structural measurements using LiDAR (Hudak, Crookston, Evans, Hall, & Falkowski, 2008) and spectral/LiDAR combinations (Guo, Chehata, Mallet, & Boukir, 2011; Leutner et al., 2012). Ensemble learning approaches such as random forests are well-suited to handle "wide-datasets" such as the datasets analyzed in this study because they result in smaller prediction variance and bias and better model performance compared to other approaches (e.g., Gislason, Benediktsson, & Sveinsson, 2006; Mitchell et al., 2013; Pal, 2005; Rodriguez-Galiano, Ghimire, Rogan, Chica-Olmo, & Rigol-Sanchez, 2012; Strobl, Malley, & Tutz, 2009).

In the first part of this study we explore the relative contributions of high resolution (3 m pixels) airborne HyMap and discrete return, small footprint LiDAR data to the estimation of shrub cover using ground reference data sampled across three collection sites that span precipitation and elevation gradients in the Great Basin region of southern Idaho, USA (Olsoy, Glenn, Clark, & Derryberry, 2014). We also consider uncertainty associated with combining all three sites for analysis. In the second part of this study, we simulate HyspIRI imaging spectrometer data to assess the potential for satellite hyperspectral data to estimate shrub cover and height at the regional scale (60 m pixels; across all three sites) using LiDAR-only metrics as a pseudo validation dataset. Findings are designed to provide insight into the extent to which hyperspectral satellite observations can augment structure measurements in dryland systems where future laser altimetry satellite technologies may be sensitive to areas of low canopy cover.

2. Methods

2.1. Study sites

The study areas consist of three collection sites located across the sagebrush-steppe ecosystem in southern Idaho, USA (Fig. 1): Department of Energy's Idaho National Lab (INL), Hollister, and Reynolds Creek Experimental Watershed (RCEW). The INL study site is located in cold desert sagebrush-steppe along the eastern Snake River plain in an intermountain landscape. The study area and its vicinity are flat, with elevations in the study area ranging from approximately 1479 to 1496 m. Microtopographical fluctuations created by historical agricultural practices, namely archaic irrigation channels and associated side channels, are present in the northeastern portion of the project area. The study site is dominated by Wyoming big sagebrush (*Artemisia tridentata* subsp. *wyomingensis*), while basin big sagebrush (*Artemisia tridentata* subsp. *tridentata*) occurs in association with depressional areas and drainage channels. Other species common to the study area include yellow rabbit brush (*Chrysothamnus viscidiflorus*), prickly pear cactus (*Opuntia* spp.) and crested wheatgrass (*Agropyron cristatum*).

The Hollister site is located in the County of Twin Falls in the Snake River plain region of southern Idaho. The study area is sloped southwest to northeast, with elevations ranging from approximately 1551 m in the southern portion to approximately 1362 m in the northern portion of the site. The plant community is Wyoming Big Sagebrush (*A. tridentata* ssp. *wyomingensis*) of low-stature (generally <50 cm, all <1 m) (Fig. 2a) and a relatively high ratio of wood: leaves. Herbaceous cover includes Sandberg's bluegrass and squirreltail (*Poa secunda* and *Elymus elemoides*, respectively) as dominant understory bunchgrasses and moderate and patchy occurrences of cheatgrass, crested wheatgrass, and native forbs. Fire history records indicate minimal disturbance.

The RCEW consists of approximately 239 km^2 of land located in the Owyhee Mountains in southwestern Idaho, USA. Elevations in the watershed range from 1049 to 2245 m. Sagebrush and grassland communities are the dominant vegetation covers (Fig. 2b). Common shrub species include low sagebrush (*Artemisia arbuscula* Nutt.), big sagebrush (*Artemisia tridentata* Nutt. subsp. *vaseyana* [Rydb.] Beetle and subsp. *wyomingensis*) and bitterbrush (*Purshia tridentata* [Pursh] DC), which typically grow up to 50 cm, 50–100 cm, and 60–185 cm in height, respectively.

2.2. Data collection

This cross-site shrub cover analysis was designed using ground reference and airborne HyMap and LiDAR data collected at three sagebrush-steppe sites in southern Idaho from 2007 to 2011 (Table 1).