Remote Sensing of Environment 210 (2018) 490–507

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Mapping of forest alliances with simulated multi-seasonal hyperspectral satellite imagery

Matthew L. Clark[a,*], Jennifer Buck-Diaz[b], Julie Evens[b]

[a] Center for Interdisciplinary Geospatial Analysis, Department of Geography, Environment, and Planning, Sonoma State University, Rohnert Park, CA 94928, USA
[b] California Native Plant Society, Sacramento, CA 95816, USA

ARTICLE INFO

Keywords:
Land cover
Forest alliance
U.S. National Vegetation Classification (NVC)
Multi-seasonal
Hyperspectral satellite
Imaging spectroscopy
HyspIRI
Machine learning
Support Vector Machine
Regional
County

ABSTRACT

A consistent and hierarchical classification of vegetation, such as the U.S. National Vegetation Classification (NVC) system, supports comprehensive conservation and management of natural ecosystems. At a detailed level, the NVC alliance is defined by diagnostic species and composition. Maps at this level of classification are often produced at local to regional scales (areas < 25,000 km^2) with costly manual to semi-automated interpretation of high resolution imagery. The main objective of this study was to assess the effectiveness of machine learning for automated, per-pixel (30 m) mapping of forest alliances with multi-seasonal hyperspectral imagery from a future satellite mission (HyspIRI), as simulated from Airborne Visible/Infrared Imaging Spectrometer Classic (AVIRIS-C) data. The study area was the San Francisco (S.F.) Bay Area, California, where we mapped forest alliances at regional and county scales. We implemented the Support Vector Machine (SVM) classifier in a two-stage approach, first mapping regional land cover followed by forest alliances in closed-canopy tree pixels. Predictor variables were reflectance bands and hyperspectral metrics based on indices, derivatives and absorption-fitting techniques applied to reflectance spectra, with data grouped into summer and three-season (spring, summer, fall) sets. For forest alliances, hyperspectral metrics improved overall accuracy of classifications by 2.9 to 6.4% relative to classifications based on the original reflectance bands. Multi-seasonal data improved overall accuracy by 1.3 to 6.2% relative to summer-only data. Using multi-seasonal metrics, the S.F. Bay Area regional classification with 21 alliances had an overall accuracy of 65.7% (Kappa 0.63), while the Sonoma County classification with 16 alliances had an accuracy of 75.9% (Kappa 0.72). Most forest alliances had internal variation in lifeform, species and structural properties that increased within-class spectral-temporal variation and complicated discrimination. Despite this challenge, classification accuracies were similar to regional NVC alliance reference data. We conclude that a hyperspectral satellite, with its repeat and global image acquisitions, has strong potential for accurate and economical mapping and monitoring the Earth's vegetation communities.

1. Introduction

An integrated and hierarchical classification of vegetation supports a comprehensive approach to conservation and management of natural ecosystems. Field data are used to define detailed floristic definitions, such as community type, while vegetation maps provide a broader view of natural and anthropogenic ecosystems and their interconnected processes across landscapes (e.g., fire, water flow, genetic exchange). Further, natural communities are a "coarse filter" delineation useful for assessing ecological health, future conservation priorities, and identifying land management risks.

The United States National Vegetation Classification (NVC) provides a detailed floristic classification scheme organized in eight hierarchical levels that identify ecologically-related vegetation (Table 1). This classification system was codified by the U.S. Federal Geographic Data Committee (FGDC, 2008) and now includes national and international levels and revised criteria for vegetation classification (Faber-Langendoen et al., 2009; Jennings et al., 2009). The NVC alliance level is defined by diagnostic plant species, mainly those from the dominant or primary growth form, and on composition that reflects regional to subregional climate, substrate, hydrology, moisture/nutrient factors, and disturbance regimes. Associations are the finest level of the NVC and describe variations within alliances based on multiple growth forms, and more narrowly similar composition. In California, the state where the current study is situated, the California Native Plant Society and California Department of Fish and Wildlife (CDFW) have

* Corresponding author.
E-mail address: matthew.clark@sonoma.edu (M.L. Clark).

https://doi.org/10.1016/j.rse.2018.03.021
Received 10 September 2017; Received in revised form 19 February 2018; Accepted 16 March 2018

Table 1
The U.S. National Vegetation Classification (NVC), with example *Quercus garryana* (Oregon white oak) Woodland Alliance (Klein et al., 2015; USNVC, 2016).

Hierarchy Level	Example Name
Level 1 - Formation Class	Forest & Woodland
Level 2 - Formation Subclass	Temperate & Boreal Forest & Woodland
Level 3 - Formation	Cool Temperate Forest & Woodland
Level 4 - Division	Vancouverian Forest & Woodland
Level 5 - Macrogroup	Southern Vancouverian Dry Foothill Forest & Woodland
Level 6 - Group	Cascadian Oregon White Oak - Conifer Forest & Woodland
Level 7 - Alliance	*Quercus garryana* (tree) Woodland Alliance
Level 8 - Association	*Quercus garryana - Umbellularia californica - Quercus* (*agrifolia*, *kelloggii*) Association

established a state classification at the alliance and association levels (Sawyer et al., 2009). Adoption of a common, standard classification has allowed for conventions, descriptions, and names to be consistent.

Floristic classifications represent species assemblages and structural details, some of which may be difficult to resolve through visual interpretation of aerial photographs or analyses of spectral signatures, particularly without supporting ground observations. A vegetation map is thus a symbolic representation of distinct groupings of plants, which is limited by the resolution of its base imagery and by its visual or spectral interpretation at a certain scale, and may not contain the finest levels of floristic classification. In the United States, mapping of NVC alliances and higher-level units is generally done at the local to regional scale (areas < 25,000 km^2) with visual interpretation of high-resolution (≤1 m) color aerial photographs by vegetation experts in conjunction with field surveys (NPS, 2013). The cost and time considerations in mapping NVC classes with high-resolution imagery are still high for creating seamless and current maps at national and broader spatial scales (CSU Northridge, 2015).

Pixel-based remote sensing using medium resolution (30 m) imagery from satellites has a role in bridging scales missed by mapping efforts covering smaller areas. However, medium resolution maps are often at higher levels in the NVC hierarchy than alliance. For example, the USGS GAP national land cover data were derived from multi-seasonal, multi-spectral satellite imagery (Landsat ETM+) and digital elevation models at 30-m pixel resolution and included vegetation categories to the NVC macrogroup level (Table 1; Wickham et al., 2014). Other studies with automated multispectral/multi-temporal approaches have explored finer levels of NVC classification, yet with much smaller spatial extents (Chastain et al., 2008; de Colstoun et al., 2003; Ramsey III et al., 2002). For example, Chastain et al. (2008) used high-resolution color-infrared imagery and multi-seasonal Landsat with discriminant analysis to map 49 NVC associations over a 330 km^2 area of Missouri. Overall accuracy was 63%, and reached 78% with 33 reduced classes.

Hyperspectral, or imaging spectrometer sensors acquire images with hundreds of narrow spectral measurements across the visible to shortwave infrared (VSWIR) that are sensitive to subtle spectral properties determined by vegetation chemistry, structure and physiology (Asner, 1998). These sensors have been used in innovative remote sensing applications focused on estimating forest canopy chemicals/plant functional traits (Asner et al., 2017; Kokaly et al., 2009) and gradients of diversity (Ceballos et al., 2015; Féret and Asner, 2014). Relative to multispectral data, hyperspectral data can improve the mapping of land-cover properties (Asner and Heidebrecht, 2002; Goodenough et al., 2003; Mariotto et al., 2013; Plourde et al., 2007; Thenkabail et al., 2004), and at high spatial resolutions (i.e., < 5 m), hyperspectral data can improve plant species discrimination (Belluco et al., 2006; Clark et al., 2005; Dalponte et al., 2012; Ferreira et al., 2016).

Largely due to lack of adequate imagery from a spaceborne sensor, most hyperspectral research has been limited to relatively small spatial extents covered by airborne sensors. The recently-decommissioned Hyperion satellite on EO-1 provided ~15 years of experimental VSWIR medium resolution (30 m) hyperspectral imagery in sample locations tasked by the community (7.7 km swath, 45–185 km scene length). These data allowed the first exploration of hyperspectral mapping of vegetation from space over larger extents yet with coarser pixel sizes than found with airborne data. Vegetation classification applications with Hyperion imagery include general land cover (Petropoulos et al., 2012; Wang et al., 2010) and hybrid schemes with dominant species (Goodenough et al., 2003; Ham et al., 2005; Pignatti et al., 2009; Vyas et al., 2011). A new generation of hyperspectral satellites could provide the global and temporal coverage necessary for a more thorough assessment of the technology over greater spatial extents and variety of ecosystems (Clark and Kilham, 2016; Jetz et al., 2016). For example, the Hyperspectral Infrared Imager (HyspIRI; Lee et al., 2015), with 210 bands, is a concept-phase NASA mission that would provide 30-m VSWIR imagery with a Landsat-style global coverage and acquisition cycle. Further, a space-based hyperspectral imager, such as HyspIRI, was recommended by the 2017–2027 Decadal Survey for Earth Science and Applications from Space as a high-priority mission (NASEM, 2018).

Research that demonstrates the capabilities and understanding of mapping community/ecosystem types, such as NVC, or dominant canopy species from the medium-resolution spaceborne hyperspectral sensors is limited. A major contribution to this line of inquiry was by Roth et al. (2015), who used a large collection of airborne hyperspectral images to test the impact of spatial resolution (20–60 m pixel size) on classification accuracy for a range of plant species and plant functional types (PFTs, similar to NVC formations) across five ecosystems in North America. This study found that the accuracy was high (83 to 99.7%) for PFT classifications and moderate to high (61.5 to 96.2%) for species classifications across sites and spatial resolutions, with a general pattern of increasing accuracy with coarser pixels (≥20 m) relative to native resolution. Species-level classification revealed that each class had its own optimal pixel resolution, depending on heterogeneity of the class (e.g., mixes of canopy species and background elements) and patch size. With higher-resolution pixels, within-class spectral variation from leaves, branches, exposed substrate and shadows can overwhelm the among-class spectral variation needed for accurate class discrimination. At coarser pixel resolutions, the authors determined that within-class variance is smoothed from spectral mixing of sub-pixel components within a larger field of view, thereby maximizing among-class variation and increasing classification accuracy (Clark et al., 2005; Roth et al., 2015). However, contrary to this finding, Schaaf et al. (2011) found that five temperate-zone PFTs decreased in accuracy when hyperspectral data were resampled from 20 to 60 m. These studies, and other studies focused on tree species classification (reviewed in Fassnacht et al., 2016), indicate that different vegetation species, classification hierarchies, and object-based delineations may have their own optimal pixel resolution. Further, reference data accuracy, species-mixture thresholds, and size/shape characteristics are important factors to consider in analyses as future research expands to include more global ecosystems (Roth et al., 2015).

Besides greater spatial coverage, hyperspectral satellites also have systematic, repeated acquisitions, providing opportunities to discriminate land-cover and species assemblages based on variation in phenology. Research with multispectral satellites has established that temporal information can aid in discriminating anthropogenic cover types (e.g., crops), monitoring decadal land change, and increasing ground observations in areas with persistent cloud cover (Clark et al., 2012; Homer et al., 2015; Zhao et al., 2016). Fagan et al., (2015) found that single-season airborne hyperspectral data (14.2 to 16.7 m) with four years of Landsat satellite data improved discrimination of general tropical land-cover/forest types over hyperspectral data alone. In our