Remote Sensing of Environment 167 (2015) 135–151

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Differentiating plant species within and across diverse ecosystems with imaging spectroscopy

Keely L. Roth [a,*], Dar A. Roberts [b], Philip E. Dennison [c], Michael Alonzo [b], Seth H. Peterson [b], Michael Beland [b,d]

[a] Department of Land, Air and Water Resources, University of California, Davis, One Shields Ave, Davis, CA 95616, United States
[b] Department of Geography, University of California, Santa Barbara, 1832 Ellison Hall, Santa Barbara, CA 93106, United States
[c] Department of Geography and Center for Natural and Technological Hazards, University of Utah, 260 S Central Campus Dr., Room 270, Salt Lake City, UT 84112, United States
[d] Department of Geography, San Diego State University, 5500 Campanile Dr., Geography Annex 123, San Diego, CA 92182, United States

ARTICLE INFO

Article history:
Received 10 June 2014
Received in revised form 10 May 2015
Accepted 15 May 2015
Available online 28 May 2015

Keywords:
Plant species classification
Hyperspectral
Imaging spectroscopy
AVIRIS
MESMA
Discriminant analysis

ABSTRACT

Imaging spectroscopy has been used successfully to map species across diverse ecosystems, and with several spaceborne imaging spectrometer missions underway (e.g., Hyperspectral Infrared Imager (HyspIRI), Environmental Mapping and Analysis Program (EnMAP)), these data may soon be available globally. Still, most studies have focused only on single ecosystems, and many different classification strategies have been used, making it difficult to assess the potential for mapping dominant species on a broader scale. Here we compare a number of classification approaches across five contrasting ecosystems containing an expansive diversity of species and plant functional types in an effort to find a robust strategy for discriminating among dominant plant species within and across ecosystems. We evaluated the performance of combinations of methods of training data selection (stratified random selection and iterative endmember selection (IES)), spectral dimension reduction methods (canonical discriminant analysis (CDA) and partial least squares regression (PLSR)) and classification algorithms (linear discriminant analysis (LDA) and Multiple Endmember Spectral Mixture Analysis (MESMA)). Accuracy was assessed using an independent validation data set. Mean kappa coefficients for all strategies ranged from 0.48 to 0.85 for each ecosystem. Maximum kappa values and overall accuracies within each ecosystem ranged from 0.56 to 0.90 and 61–92%, respectively. Our findings show that both LDA and MESMA are able to discriminate among species to a high degree of accuracy in most ecosystems, with LDA performing slightly better. Spectral dimension reduction generally improved these results, particularly in conjunction with MESMA. Within each ecosystem, both the number and identities of functional types present, as well as the spatial distribution of dominant species, played a strong role in classification accuracy. In a pooled ecosystem classification, using CDA and LDA, we discriminated among 65 classes with an overall accuracy of 70% for the validation library, using only a 6% training sample. Our results suggest that a spaceborne imaging spectrometer such as HyspIRI will be able to map dominant plant species on a broader scale.

1. Introduction

Accurate information regarding the composition and distribution of dominant plant species and, therefore, plant functional types, within and across ecosystems is pertinent to many research agendas within ecosystem science and plant ecology. Species maps allow scientists to detect the presence or absence of target species (e.g., invasive species, He, Rocchini, Neteler, & Nagendra, 2011; Somers & Asner, 2012) and monitor landscape-scale biological changes such as distribution shifts (Asner, Jones, Martin, Knapp, & Hughes, 2008), type conversion, and disturbance impact and recovery (Hatala, Crabtree, Halligan, & Moorcroft, 2010; Riano et al., 2002). This information is also critical for further refining estimates of ecosystem function (e.g., biomass, habitat suitability), and thus providing improved spatially explicit inputs for evolving ecosystem process and climate models (Goodenough et al., 2006; Kokaly, Asner, Ollinger, Martin, & Wessman, 2009; Ustin, Roberts, Gamon, Asner, & Green, 2004).

Improvements in sensor technology and the development of more sophisticated classification algorithms have enabled remote sensing scientists to discriminate among various vegetation communities (e.g., forest, crop, grassland) and life forms (e.g., herbaceous, shrubs, trees) (DeFries, Hansen, & Townshend, 1995; Friedl et al., 2010), between different leaf types (i.e., broadleaf vs. coniferous) (Van Aardt & Wynne, 2001) and among plant functional types (PFTs) (e.g., deciduous broadleaf tree, evergreen needleleaf shrub) (reviewed in Ustin & Gamon, 2010). However, discriminating individual plant species requires data with fine spectral resolution, which can be acquired using imaging spectrometers (Clark,

* Corresponding author.
E-mail address: klroth@ucdavis.edu (K.L. Roth).

http://dx.doi.org/10.1016/j.rse.2015.05.007
0034-4257/

Roberts, & Clark, 2005). Imaging spectrometers are sensitive to subtle shifts in spectral properties that are controlled by leaf biochemistry, anatomy and physiology and are further modified by canopy architecture (Asner, 1998; Roberts et al., 2004). As such, these instruments have been used successfully to discriminate among plant species and functional types using leaf-level observations (Castro-Esau, Sanchez-Azofeifa, & Caelli, 2004; Clark et al., 2005), field-collected canopy spectra (Gong, Pu, & Yu, 1997; Pu, 2009), and image data (Martin, Newman, Aber, & Congalton, 1998; Van Aardt & Wynne, 2007). Indeed, a major aim of several upcoming global hyperspectral missions is to map plant species and functional types in support of ecosystem research, including NASA's Hyperspectral Infrared Imager (HyspIRI; HyspIRI Team, 2009). HyspIRI would collect full Visible/Near-Infrared/Short-wave infrared (VNIR-SWIR) spectra (400–2500 nm) at 60 m spatial resolution on a global, 19 day repeat cycle.

The ability to discriminate accurately among dominant plant species and PFTs on regional to global scales represents a major advance in remote sensing science (Asner, 2013). However, success depends, in large part, on a solid understanding of the spectral, spatial and temporal resolution constraints on mapping species within and across a diverse set of ecosystems. Indeed, most imaging spectroscopy studies have sought to measure the spectral separability of, or to classify, species in single ecosystems or species in single plant functional types (e.g., Féret & Asner, 2012a; Kalacska, Bohlman, Sanchez-Azofeifa, Castro-Esau, & Caelli, 2007; Pu, 2009; Vaiphasa, Ongsomwang, Vaiphasa, & Skidmore, 2005; Van Aardt & Wynne, 2007) and thus have investigated a limited diversity of species and functional types, making comparisons across ecosystems challenging. Furthermore, most studies that explicitly evaluate the effects of dimension reduction (Dópido et al., 2012; Kalacska et al., 2007; Pu & Gong, 2000) or classification method (Clark et al., 2005; Féret & Asner, 2012b) on classification accuracy, have performed such analyses on only a single test data set or within a single ecosystem, or do not focus on discriminating among plant species. None, to our knowledge, evaluate the impact of different combinations of methods across a diversity of ecosystems.

To fully leverage the data provided by a global imaging spectrometer, such as HyspIRI, we must improve our understanding of the spectral properties of a diverse range of species and PFTs in the landscapes we seek to map, the methods we use to create these maps and how the two interrelate. Thus, the goals of this research were to evaluate our ability to spectrally discriminate dominant plant species in contrasting ecosystems and to compare the performance of several hyperspectral classification strategies in accurately mapping species across multiple, diverse ecosystems. Our main research questions are as follows:

1) Within individual ecosystems, how spectrally separable are the dominant species and what ecosystem characteristics drive observed separability?
2) When applied to a diverse set of ecosystems, how do classification strategies differ in performance, i.e. is there a clearly superior strategy?
3) What is the potential for differentiating among species from all ecosystems using the best approach as determined by question 2?

Comparing different classification approaches across ecosystems, we can evaluate if (and how) the methods perform differently for different ecosystems. This will indicate if there is a best overall approach, or if different approaches are needed depending on the ecosystem. By applying the same classification methods at each site, we highlight our ability to spectrally separate species in each ecosystem type. In other words, we can characterize what makes one ecosystem easier to map vs. another, and explore the possibility of a general limit to how accurately dominant species within a particular ecosystem type can be classified with imaging spectroscopy data alone. By combining the ecosystems, we are testing our ability to map species across multiple ecosystems simultaneously, which will be the goal for the larger footprint spaceborne hyperspectral data collected by a sensor like HyspIRI. Does such a classification yield acceptable results? How are class-level accuracies affected (e.g., are some species classified more accurately with higher diversity in the classification? are species within the same plant functional type but from different ecosystems spectrally confused?)? Do we see similar patterns in misclassification when all sites are grouped together as we do when we map sites individually? The answers to these questions can provide great insight for future large scale species mapping efforts.

2. Methods

2.1. Study sites & data collection

We analyzed image data from five different North American ecosystems (Fig. 1, Table 1). The Smithsonian Environmental Research Center (SERC) site is a temperate, broadleaf deciduous forest in eastern Maryland ranging in elevation from 2 to 20 m. It is dominated by intermediate to mature stands of tulip poplar (*Liriodendron tulipifera*) and sweetgum (*Liquidambar styraciflua*) mixed with maple (*Acer* spp.), hickory (*Carya* spp.) and beech (*Fagus* spp.). The forested area of the site is surrounded primarily by agriculture and open fields. The Gulf study site is located in coastal Louisiana, including Barataria Bay and the Mississippi delta, with elevations just above sea level. It is a marsh ecosystem strongly influenced by a subtropical climate and the confluence of fresh and salt water. Cordgrass (*Spartina* spp.), Salt grass (*Distichlis spicata*), and black rush (*Juncus roemerianus*) dominate the salt and brackish marsh zones, and common reed (*Phragmites australis*) becomes prevalent in intermediate to fresh water zones. The Wind River Experimental Forest (WR) site is a mixed broadleaf and coniferous temperate rainforest located in southern Washington in the Cascade Mountains, covering an elevation gradient of approximately 250 to 800 m. It is dominated primarily by western hemlock (*Tsuga heterophylla*) and Douglas fir (*Pseudotsuga menziesii*), with an herbaceous understory and smaller stands of maple (*Acer* spp.), cottonwood (*Populus trichocarpa*) and alder (*Alnus rubra*). The Sierra Nevada site (SNEV) is a mixed montane coniferous forest in the southern Sierra Nevada Mountains of California. The site includes major portions of the Sierra National Forest, extending from Shaver Lake southeast to Kings Canyon National Park and covering a range in elevation from approximately 1200 to 2000 m. It is composed of large mixed stands of fir (both white and red, *Abies concolor* and *magnifica*, respectively) and pine (Ponderosa, Jeffrey and sugar; *Pinus ponderosa*, *jeffreyi*, and *lambertiana*, respectively), as well as broad swaths of deciduous and evergreen oak (*Quercus kelloggii* and *Quercus chrysolepsis*, respectively) with shrub-dominated rocky outcrops, open meadows and riparian zones. The Santa Barbara (SBFR) site runs east to west along the front range of the coastal Santa Ynez Mountains in southern California, and extends north from Santa Barbara into the San Raphael Mountains. It covers a large swath of shrublands, grasslands, woodlands and urban areas distributed over 1 to 1366 m elevation, and has a Mediterranean climate, with cool, moist winters and dry, warm summers. Wooded regions are dominated by oak (*Quercus* spp.) and California bay laurel (*Umbellularia californica*), with some patches of sycamore (*Platanus racemosa*) and gray pine (*Pinus sabiniana*). Major chaparral shrub species include *Ceanothus* spp., chamise (*Adenostoma fasciculatum*), and manzanita (*Arctostaphylos* spp.).

Image data were acquired by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) over the five study regions (Table 1). AVIRIS collects data in 224 bands from 350 to 2500 nm with a full-width, half-maximum of about 10 nm (Green et al., 1998). Data were preprocessed to radiance and georectified using a ray tracing algorithm with a digital elevation model (Boardman, 1999). Reflectance was retrieved for all images using either MODTRAN-derived look-up tables for path and reflected radiance (described in Roberts, Green, & Adams, 1997), ACORN (ImSpec LLC) or ATCOR-4 (Richter & Schläpfer, 2002). Bands with low signal to noise ratio and/or high levels of atmospheric