Remote Sensing of Environment 168 (2015) 286–300

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

# Characterization of shrubland ecosystem components as continuous fields in the northwest United States

George Xian [a,*], Collin Homer [a], Matthew Rigge [b], Hua Shi [b], Debbie Meyer [c]

[a] U.S. Geological Survey (USGS) Earth Resources Observation and Science Center, Sioux Falls, SD 57198, USA
[b] InuTeq/U.S. Geological Survey (USGS) Earth Resources Observation and Science Center, Sioux Falls, SD 57198, USA
[c] SGT/USGS Earth Resources Observation and Science Center, Sioux Falls, SD 57198, USA

ARTICLE INFO

*Article history:*
Received 31 October 2014
Received in revised form 5 July 2015
Accepted 16 July 2015
Available online 31 July 2015

*Keywords:*
Shrubland
Ecosystems
Landsat 8
Image mosaic
Climate

ABSTRACT

Accurate and consistent estimates of shrubland ecosystem components are crucial to a better understanding of ecosystem conditions in arid and semiarid lands. An innovative approach was developed by integrating multiple sources of information to quantify shrubland components as continuous field products within the National Land Cover Database (NLCD). The approach consists of several procedures including field sample collections, high-resolution mapping of shrubland components using WorldView-2 imagery and regression tree models, Landsat 8 radiometric balancing and phenological mosaicking, medium resolution estimates of shrubland components following different climate zones using Landsat 8 phenological mosaics and regression tree models, and product validation. Fractional covers of nine shrubland components were estimated: annual herbaceous, bare ground, big sagebrush, herbaceous, litter, sagebrush, shrub, sagebrush height, and shrub height. Our study area included the footprint of six Landsat 8 scenes in the northwestern United States. Results show that most components have relatively significant correlations with validation data, have small normalized root mean square errors, and correspond well with expected ecological gradients. While some uncertainties remain with height estimates, the model formulated in this study provides a cross-validated, unbiased, and cost effective approach to quantify shrubland components at a regional scale and advances knowledge of horizontal and vertical variability of these components.

Published by Elsevier Inc.

## 1. Introduction

Shrubland ecosystems cover a large portion of the global land surface, mostly in semiarid ecosystems. For example, several global land cover datasets have estimated that shrub cover occupies about 14% (Broxton et al., 2014; Loveland et al., 2000) to 17% (Hansen et al., 2000) of the global land. Shrub cover was consistently estimated as 21% of total land across the conterminous United States in 2001 and 2006 (Fry et al., 2011). Shrubland ecosystems are critical for determining regional biodiversity, global biogeochemical cycles, and energy and gas fluxes (Hamada et al., 2013; Schwinning et al., 2005; Wallace et al., 2003). Shrublands provide many ecosystem services to human populations around the world (Sant et al., 2014). Overall, shrub and grass cover about 2.9 million km$^2$, or 36% of the total land in the conterminous United States (Fry et al., 2011). In western North America, shrubland and grassland comprise the largest rangeland vegetation biomes. Sagebrush (*Artemisia* spp.) communities alone are estimated to cover 0.23 million (Branson et al., 1967) to 0.25 million km$^2$ (West, 1983) in the western United States. The sagebrush area is about 13.3 to 14.5% of the total shrubland in the conterminous Unites States.

Many recent changes, including diminishing vegetation cover, encroachment of woody and invasive species, and the degradation of shrub and grass communities in general, have been observed in arid and semiarid ecosystems throughout the Western United States where much of the country's shrubland community occurs (Connelly et al., 2004; Schroeder et al., 2004; Bradley et al., 2006; Homer et al., 2009; Homer et al., 2012; Xian et al., 2012a). For example, the most dominant shrub community in the Western United States, sagebrush, has suffered steadily declining quality and quantity from the historical distribution. In recent decades, the speed of the deterioration has accelerated due to the effects and interactions among land use change, livestock grazing, fire, invasive species, and climate change (Berlow et al., 2002; Davies et al., 2007; Olsoy et al., 2014). Recent natural gas extraction in western Wyoming has further significantly affected sagebrush communities in the region by reducing and degrading sagebrush habitat (Walston et al., 2009). Cheatgrass (*Bromus tectorum*), an invasive species, and wildland fire also reduced the abundance of sagebrush shrublands (Bradley et al., 2006; Knick et al., 2003). Climate presents another important challenge to these ecosystems (Brown et al., 1997; Munson et al., 2011). Most shrubland components in semiarid lands are vulnerable to climate variations, especially to precipitation changes, because precipitation effects on

* Corresponding author.
*E-mail address:* xian@usgs.gov (G. Xian).

http://dx.doi.org/10.1016/j.rse.2015.07.014
0034-4257/Published by Elsevier Inc.

plants are mediated primarily by the distribution and dynamics of soil water interacting with species' roots, which usually exist in low soil moisture content (Reynolds, Kemp & Tenhunen, 2000; Schwinning et al., 2005; Weltzin et al., 2003). Furthermore, variations in precipitation strongly influence arid and semiarid land plant composition and dynamics (Branson et al., 1976; Cook & Irwin, 1992; Pelaez et al., 1994; Reynolds et al., 2000; Xian et al., 2012b), even at higher elevations (Guida et al., 2014). The long-term climate data analysis shows that for the period 1895–2011 in the northwest region, the most recent 35 years have included a few very wet years, including the wettest year on record in 1996 (Kunkel et al., 2013). However, the majority of the recent years have been below the 1901–1960 average, including the multi-year dry periods of 2007–2009 and 2000–2002. General precipitation trends in these regions have not shown statistically significant trends for any season. However, temperatures have generally been above the long-term average for the last 25 years, both annually and for all four seasons.

Increasingly, the decline in the spatial extent and deterioration in the quality of shrubland ecosystem components reduces the ability of shrubland to provide the same level of ecosystem services in arid and semiarid lands. Improved characterization and monitoring of shrubland ecosystem components are needed to facilitate natural resources management against more complex future dynamics (Hemstrom et al., 2002). Furthermore, the information can be used to assess and monitor changes in shrubland through time and across large geographic areas. However, for highly heterogeneous shrubland ecosystems, vegetation cover usually has a large range of variability from land dominated by bare ground to areas of high density shrubs. Also, it can be difficult to quantify true cover amounts of different components using only thematic land cover characterization because highly heterogeneous vegetation components such as herbaceous, sagebrush, and shrub usually occur in close approximation. This results in large uncertainties for biomass estimates and change assessments. For better addressing the spatial distributions of shrubland components, these components need to be characterized as continuous fields (Xian et al., 2013).

Traditional field-based methods used to assess and monitor spatial distributions of shrubland components have limitations in accuracy, efficiency, and consistency (Olsen et al., 1999). Field observations are generally confined to short-term field measurements with limited samples from limited sampling sites. Given their vast geographic extents and complicated nature of shrubland ecosystems, remote sensing data that have large area coverage and frequent acquisition cycle are often preferred to inventory, monitor, and characterize shrublands. These efforts include the use of high-resolution remotely sensed data such as QuickBird imagery (Laliberte et al., 2004), aerial photography (McGlynn & Okin, 2006), ground-based natural color vertical photography (Stow et al., 2008), LIDAR data (Mitchell et al., 2011; Sankey & Bond, 2011), and radar data (Huang et al., 2010). There are trade-offs with remotely sensed data between resolution and the extent of area imaged, with high resolution data capturing small extents and low resolution data capturing large extents. Usually, high-resolution imagery can characterize shrubland components with higher accuracy but represent small extents. Other applications have adopted medium resolution Landsat data (Hamada et al., 2011; Hamada et al., 2013; Homer et al., 2009; Ramsey et al., 2004; Sant et al., 2014; Sivanpillai & Ewers, 2013; Sivanpillai et al., 2009; Viedma & Medliá, 1999; Vogelmann et al., 2012; Wallace et al., 2003; Xian et al., 2012a) or MODIS imagery (Wallace et al., 2008) to assess spatial cover and distribution of shrubland components at a relatively large scale. The recent developments in estimating shrubland components as continuous fields using medium resolution Landsat imagery and regression tree (RT) models have been effectively demonstrated for large-scale areas (Homer et al., 2012; Xian et al., 2013). However, these approaches were restricted to functioning with one individual Landsat path/row at a time and limited the potential to scale up to large areas that are extending beyond Landsat path/row extents. In order to operationally cover larger geographies, it is necessary to find approaches that can handle multiple Landsat scenes for a much larger geospatial extent than a single Landsat path/row.

Generally, mosaicking images from multiple path/rows can improve the efficiency of a large area mapping effort by optimizing ground training data and allowing classification to be performed on the whole mosaic for a large area. However, phenological differences between scenes of different dates affect the quality and accuracy of the modeling when the common mosaicking approaches are used for RT model products.

To reduce phenological effects for image change detection, different normalization approaches, including spectral information of an entire Landsat image (Xian et al., 2009; Yang & Lo, 2000) and relative radiometric normalization (Du et al., 2001; Elvidge et al., 1995; Jensen, 1996), have been attempted to match image radiometric signals so that they are consistent between images. However, the scene-based linear regression normalization usually fails to represent the complexity of pixel-level differences for images acquired from neighboring paths/rows with different dates. Relative radiometric normalization, especially overlap regression normalization methods, have been shown to produce consistent radiances in overlap regions between scenes to produce a visually seamless mosaic. Theil–Sen regression, which provides unbiased regression line fits in the presence of measurement errors in both variables in contrast to classical linear regression approaches, has been implemented for Landsat ETM + normalization (Olthof et al., 2005). The method uses the median of all slopes of unique pairwise observations as the maximum likelihood slope for the regression and is robust without performing procedures to detect or remove outliers prior to regression. However, the Theil–Sen regression method heavily relies on the median slope to form simple linear regression equations. The median slope-based regression equations might not effectively represent complicated landscape conditions for shrubland ecosystems and phenological differences in the normalized mosaic could result in biases for the continuous field quantification. Different radiometric balancing methods are required in order to characterize shrubland ecosystem components for a large geographic extent with Landsat data in a more robust manner.

Another challenge of estimating shrubland component distributions at a large scale using remote sensing data and modeling is how to incorporate external constraints such as climate and ecological conditions to improve model estimates. The large spatial extent covered by several Landsat images in the western United States, for example, usually contains several climate zones (Peel et al., 2007). Populations of shrubland ecosystems persist within a range of climatic conditions known as their climate envelope. The climatic regime can serve as a sensitivity proxy for revealing the variability and the spatial distribution of shrubland ecosystems that are primarily controlled by local thermal and moisture regimes. It is therefore often necessary to characterize climate features into different regions where shrubland ecosystems are situated within a relatively homogeneous climate regime. However, the RT models are usually influenced by dominant cover ranges of a target component that might exist in one or two climate or ecological zones. The modeling result, therefore, might not be able to effectively represent the spatial distributions of different shrubland components, especially when components are non-uniformly distributed and extends across several ecological regions. Including climate condition in large-scale shrubland component modeling procedure may be key to locally stratifying shrubland ecosystem components successfully across large geographic extents.

Our overall goal is to explore a method that can integrate both field samples and remote sensing data to characterize shrubland ecosystem components as continuous fields in a large geographic extent. In this study, we focus on several specific objectives. First, we hypothesize that phenological effects of adjacent Landsat images can be reduced by using a multi-variant regression approach. Newly released Landsat 8 images acquired from different paths/rows on different dates are used to generate seasonal seamless phenological mosaics for the entire