Remote Sensing of Environment 129 (2013) 90–102

Contents lists available at SciVerse ScienceDirect

Remote Sensing of Environment

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

Estimating the fractional cover of growth forms and bare surface in savannas. A multi-resolution approach based on regression tree ensembles

Ursula Gessner [a,b,*,1], Miriam Machwitz [a,c], Christopher Conrad [a], Stefan Dech [a,b]

[a] Department of Remote Sensing, University of Wuerzburg, Am Hubland, 97074 Wuerzburg, Germany
[b] German Aerospace Center (DLR), German Remote Sensing Data Center (DFD), Muenchner Strasse 20, 82234 Oberpfaffenhofen, Germany
[c] Centre de Recherche Public — Gabriel Lippmann, Department of Environment and Agro-Biotechnologies (EVA), 41, rue du Brill, 4422 Belvaux, Luxembourg

ARTICLE INFO

Article history:
Received 13 February 2012
Received in revised form 25 September 2012
Accepted 12 October 2012
Available online 28 November 2012

Keywords:
Sub-pixel fractional cover
Savanna
Grassland and shrubland biome
Africa
Vegetation structure
Land cover
Multi-resolution analysis

ABSTRACT

Evaluations of existing land cover maps have revealed that high landscape heterogeneity and small patch sizes are a major reason for misclassification. These problems globally occur in biomes of mixed vegetation structure and are particularly relevant for African savannas. This paper presents a multi-resolution approach to derive fractional cover of vegetation growth forms at sub-pixel level, aiming at an improved mapping of land cover in the African grassland, savanna and shrubland biome. Fractional cover is delineated for woody growth forms (trees and shrubs), herbaceous growth forms, and bare surface. The approach incorporates very high resolution (QuickBird/IKONOS, 0.6–1 m), high resolution (Landsat TM/ETM+, 30 m), and medium resolution data (MODIS, 250 m). While QuickBird/IKONOS data are classified into discrete classes, at Landsat and MODIS resolutions, sub-pixel cover is delineated using non-parametric ensemble regression trees from the random forest family. The propagation of errors in the hierarchical multi-resolution approach is assessed with Monte Carlos simulations.

The multi-resolution approach allows the adequate description of the heterogeneous vegetation structure in selected study regions of Southern Africa. The RMSE of the delineated fractional cover values range between 3.1% and 8.2% when compared with higher resolution data and between 4.4% and 9.9% when compared with field surveys. Errors at the Landsat resolution show minor influence on the accuracy of the MODIS results. Regarding the inter-resolution error propagation, for 90% of the Monte Carlo simulations, errors at the Landsat resolution resulted in RMSEs for MODIS increased by less than 4% (woody vegetation), 3.5% (herbaceous vegetation) and 2% (bare surface).

1. Introduction

Information on land cover and land cover change is fundamental in many fields of environmental research. Land cover datasets are an important basis for monitoring processes in the context of global change such as deforestation, degradation, and urban expansion. Moreover, they are required in fields such as biodiversity, water management, carbon storage, and ecosystem functions. The Global Terrestrial Observing System (GTOS) selected land cover as one of the Terrestrial Essential Climate Variables, i.e. as a terrestrial property which controls physical, biological and chemical processes with relevance to climate (GTOS, 2008).

Accuracy assessments and comparisons between land cover maps have proven their general value. But likewise, studies have revealed that best accuracies are found for homogeneous land cover classes while heterogeneous classes with a mixed vegetation structure such as savannas are not as well represented. Small patch sizes and high heterogeneity were identified to be major reasons for misclassifications (Latifovic & Olthof, 2004; Smith et al., 2002, 2003). The validation of the global datasets GlobCover and MODIS Land Cover Type (MCD12Q1) revealed elevated confusions for mixed and spatially heterogeneous classes (Bicheron et al., 2008; Bontemps et al., 2011; Friedl et al., 2010). In a comparative study of two global land cover maps, Giri et al. (2005) found major differences in class assignment for heterogeneous classes. Likewise, Herold et al. (2008) identified a correlation between the homogeneity of a landscape and the agreement among four global land cover datasets. They identified low accuracies in all datasets particularly for classes with mixed vegetation such as ‘mixed trees’, ‘shrublands’ and ‘herbaceous vegetation’. Herold et al. (2008) see

* Corresponding author at: DLR-DFD, Muenchner Strasse 20, 82234 Oberpfaffenhofen, Germany. Tel.: +49 8153 28 1250; fax: +49 8153 28 1445.
E-mail addresses: ursula.gessner@dlr.de (U. Gessner), miriam.machwitz@lippmann.lu (M. Machwitz), christopher.conrad@uni-wuerzburg.de (C. Conrad), stefan.dech@dlr.de (S. Dech).
[1] Ursula Gessner did this work at the Department of Remote Sensing at the University of Wuerzburg. Her present affiliation and contact address is at DLR-DFD. Miriam Machwitz also did this work at the Department of Remote Sensing at the University of Wuerzburg. Her present affiliation and contact address is at Centre de Recherche Public — Gabriel Lippmann.

0034-4257/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.rse.2012.10.026

an outstanding challenge for improved global land cover derivation in enhancing the mapping of heterogeneous landscapes. For the African continent, these mixed classes are mainly found in the tropical and subtropical grassland, savanna and shrubland biome (biome definition according to Olson et al., 2001).

The mentioned inconsistencies may arise from different causes. Evidently, an important fact is that current land cover mapping activities rely on the distinction of discrete land cover classes. In biomes such as the tropical and subtropical savanna, grassland and shrubland biome, land cover is not discrete but rather characterized by a mixture of growth forms, primarily grasses, shrubs and trees. Small patches and gradual transitions between open and closed vegetation cover are typical features of these landscapes while clear boundaries between distinct land cover classes are rare.

A better representation of heterogeneous landscapes could be achieved by mapping the fractional cover of relevant land cover elements on a continuous scale, a concept that has been suggested by DeFries et al. (1995), Hansen et al. (2003), Heiskanen (2008) and Herold et al. (2008). Datasets following this concept of continuous categories have already been derived from medium resolution remote sensing data. The first maps of this kind focused exclusively on the derivation of tree cover. Examples are the global map of percentage tree cover, produced by DeFries et al. (2000) from AVHRR data of 1992/93 at a spatial resolution of 1 km, and a regional mapping approach for tropical woodlands and parklands in Zambia based on MODIS data (Hansen et al., 2002b, 2005). Another regional approach for forest biomes in southeast Asia was presented by Tottrup et al. (2007), considering even different forest types. A dataset that differentiates fractional cover of more than just one growth form is the MODIS Vegetation Continuous Fields product (VCF, MOD44B; Hansen et al., 2002a, 2003). Two editions of this product are currently available. The collection 3 dataset gives information on the sub-pixel proportion of trees, herbaceous vegetation and bare area for the year 2001 at a resolution of 500 m. The MODIS-VCF collection 5 contains proportional estimates only for tree cover and for the years 2000 to 2010 at a resolution of 250 m. The MODIS-VCF dataset has proven valuable especially in areas of high tree cover for mapping percent tree cover in Zambia (Hansen et al., 2005) and for depicting general tendencies of forest change in southeast Asia (Tottrup et al., 2007). However, regional assessments showed that pixel-level MODIS-VCF information has clear limitations for the monitoring of tree cover (White et al., 2005), particularly in landscapes of low tree density (Hansen et al., 2005; Heiskanen, 2008; Montesano et al., 2009). Hansen et al. (2005) state that the global land cover map that was used as a training dataset for MODIS-VCF exhibits particularly low accuracies for wooded grassland and woodland classes (i.e. savanna classes) which are due to high within-class heterogeneity. Thus, fractional vegetation cover mapping needs enhancement, especially in regions of mixed vegetation structure, where discrete land cover maps also exhibit inconsistencies (e.g. Bicheron et al., 2008; Friedl et al., 2010; Herold et al., 2008).

Apart from qualitative considerations, definitional problems arise when using tree cover or MODIS-VCF datasets in the African grassland, savanna and shrubland biome. These datasets account only for trees and partly (in MODIS-VCF, collection 3) delineate non-tree vegetation as a whole, i.e. herbaceous and shrub vegetation cover as one layer. However, in this biome, the differentiation between woody (trees and shrubs) and herbaceous growth forms seems more adequate as the ecosystem services (ESS) of shrubs and trees are generally similar, while grasses show rather different characteristics. Woody and herbaceous components exhibit for example clear differences in magnitude and seasonality of carbon sequestration, in provision of fodder and food, in the suitability of ecosystems as habitats, in their influence on the water cycle, water quality, and soil nutrients (e.g. Eldridge et al., 2011; Lloyd et al., 2008; Meik et al., 2002; Scholes & Archer, 1997).

With the aim to improve mapping of land cover in heterogeneous landscapes, this paper applies a multi-resolution approach for deriving fractional vegetation cover at regional scale. The procedure is tailored to the African grassland, savanna and shrubland biome. Thus, fractional cover is derived for the three most prevalent – and with respect to ESS most meaningful – elements of land cover in this biome: woody growth forms (trees and shrubs), herbaceous growth forms and bare surface. The approach operates on three spatial resolutions incorporating very high resolution imagery (QuickBird/IKONOS, 0.6–1 m), multi-temporal high resolution data (Landsat TM/ETM+, 30 m) and medium resolution, annual time series (MODIS, 250 m). Sub-pixel proportions of growth forms are considered at Landsat and MODIS resolution. The fractional cover derivation is based on non-parametric ensemble regression trees from the random forest family. The approach is applied for two Southern African savanna ecosystems, the Kalahari Woodland Savanna and the Central and Eastern Namibian Savanna. The propagation of errors during the hierarchical multi-resolution approach is considered using Monte Carlos simulations.

2. Study regions

The first study region, the Kalahari Woodland Savanna, is located in northeastern Namibia (Fig. 1). The area is covered by Kalahari sands with longitudinal dunes and only minor variations in elevation (1050–1250 m above sea level). Typical for the region are deep, sandy soils with low nutrient contents (ferralic Arenosols) and associated petric Calcisols (FAO, 1998; Strohbach & Petersen, 2007). The mean annual precipitation ranges between 450 mm in the southwest and 750 mm in the northeast (Mendelsohn et al., 2002). In winter, temperatures can fall below 8 °C while summer temperatures rise up to 32–36 °C (Mendelsohn et al., 2002). The vegetation is characterized by a heterogeneous mixture of open woodland savanna (*Burkeo-Pterocarpetea*) and closed to open shrublands (*Acacietea*; Strohbach & Petersen, 2007). The natural savanna vegetation is considerably influenced by human activities as the rural population strongly depends on natural resources for their livelihoods. In the Kalahari Woodland Savanna, grazing and cropping are typical land use activities, and large hardwood species are logged for timber. Frequent and mostly land-use-related savanna fires are typical and lead to a decrease in shrub density (Strohbach & Petersen, 2007).

The second study region, the Central and Eastern Namibian Savanna (Fig. 1), comprises the vegetation types 'Camelthorn Savanna' and 'Thornbush Shrubland' (Giess, 1998). The elevation above sea level varies between 1200 m and 1700 m. The most important soils in this region are ferralic Arenosols on Kalahari sands with associated petric Calcisols as well as chromic Cambisols and Regosols on sandstones and schists (Mendelsohn et al., 2002). Annual rainfall ranges from 250 mm in the southeast to 500 mm in the northeast. In winter, temperatures can drop to 0 °C while maximum summer temperatures range between 30 and 36 °C (Mendelsohn et al., 2002). The vegetation of the study area is dominated by fine-leaved and broad-leaved shrub- and grassland savannas. Fine-leaved *Acaciateae* can be found on relatively nutrient-rich soils while broad-leaved *Combretaceae* grow on poorer substrates (Rothauge, 2006). Towards the north of the region, *Terminalia–Combretum* associations are common (Strohbach et al., 2004). Land use in the Central and Eastern Namibian Savanna is mainly characterized by livestock farming on large tracts. Farmers suppress fires as far as possible and thus the savanna vegetation burns only exceptionally in this region.

Both study areas show a mixed composition of vegetation growth forms with gradual transitions between more open and more closed areas, which is a typical feature of semi-natural savanna ecosystems. Fig. 2 illustrates this characteristic savanna vegetation structure in the Kalahari Woodland Savanna. The scatter plot is based on a QuickBird classification and displays the sub-pixel proportions of growth forms within 30 m × 30 m grid cells. It is obvious that already at this spatial resolution, the savanna landscape is dominated by a heterogeneous mixture of herbaceous and woody vegetation with