Remote Sensing of Environment xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Mapping land cover in complex Mediterranean landscapes using Landsat: Improved classification accuracies from integrating multi-seasonal and synthetic imagery

Cornelius Senf *, Pedro J. Leitão, Dirk Pflugmacher, Sebastian van der Linden, Patrick Hostert

Geography Department, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany

ARTICLE INFO

Article history:
Received 20 December 2013
Received in revised form 10 October 2014
Accepted 13 October 2014
Available online xxxx

Keywords:
Landsat
Mediterranean
Pseudo-steppe
STARFM
Phenology
Image classification

ABSTRACT

Low-intensity farming systems are of great importance for biodiversity in Europe, but they are often affected by soil degradation or economic pressure, leading to either land abandonment or intensification of agriculture. These changes in land use influence local biodiversity patterns and require annual monitoring of land cover. To accurately map land cover in such spatio-temporal complex landscapes, it is important to capture their phenological dynamics and fine spatial heterogeneity. Multi-seasonal analyses using optical sensors with a medium spatial resolution from 10 to 60 m (e.g. Landsat) have been used for this task, but data availability can be scarce due to cloud cover, sub-optimal acquisition schedules and data archive access restrictions. Combining coarse spatial resolution data from the MODerate-resolution Imaging Spectroradiometer (MODIS) and Landsat provides opportunities to close these gaps by simulating Landsat-like images at MODIS temporal resolution. In this study, we test whether and by what degree land cover maps of complex Mediterranean landscapes improve by integrating multi-seasonal Landsat imagery, as well as whether STARFM-simulated imagery can be used whenever original multi-seasonal Landsat observations are unavailable. Therefore, we develop different classification scenarios based on seasonally varying data availability and based on original and simulated Landsat data. Results show that multi-seasonal Landsat data from spring and early autumn are crucial for achieving satisfying mapping accuracies (overall accuracy 74.5%). Using synthetic Landsat imagery increases classification accuracy compared to using single-date Landsat data, but accuracies were never as good as a classification based on original data. We conclude that multi-seasonal data is essential for mapping complex Mediterranean landscapes and that STARFM can be used to compensate for missing Landsat observations. However, if Landsat data availability is sufficient to cover all phenologically important dates, we suggest relying solely on Landsat.

1. Introduction

More than 50% of Europe's most valuable biotopes occur on low-intensity farming systems (Bignal & McCracken, 1996). The Iberian pseudo-steppe landscapes are one example of such low-intensity farming systems created by extensive management with long fallow period rotation systems, and are characterized by complex mosaics of (grazed) grasslands, cereal fields (mostly wheat and oat), plowed fields, patches of open cork and holm oak (*Quercus suber* and *Quercus ilex*) woodlands ("Montado") and shrubs such as rockrose, i.e. *Cistus* spp. (Leitão, Moreira, & Osborne, 2010). Pseudo-steppe landscapes are particularly important for bird biodiversity because they hold significant populations of often-protected steppe bird species (Delgado & Moreira, 2000; Moreira et al., 2007). However, due to economic pressure and generally low soil fertility, many pseudo-steppe landscapes have been either intensified (through monocultures and irrigation) or abandoned, leading to substantial changes in the local habitat availability (Moreira et al., 2012). Because these changes have direct impacts on the local pseudo-steppe bird populations (Moreira et al., 2012), mapping and monitoring the remaining pseudo-steppe habitats are required for conservation planning.

Species habitats are often strongly associated with land cover, which is one of the most prominent products derived from remote sensing (Laurent et al., 2005; McDermid, Franklin, & LeDrew, 2005). Studies have identified several land cover categories that best represent the habitat guilds of pseudo-steppe bird communities (Leitão et al., 2010; Moreira et al., 2007). Of particular importance is the occurrence of plowed fields, cereal fields, grassland, areas dominantly covered by shrubs and areas dominantly covered by trees. Many of these land cover categories have similar spectral characteristics at certain points in time (e.g. cereal fields and grassland in summer throughout the winter, or shrub- and woodlands in spring; see Fig. 3), but vary throughout the year (Guo, Price, & Stiles, 2003; Leitão et al., 2010; Toivonen & Luoto, 2003). Therefore, mapping these dynamic landscapes is best accomplished using multi-seasonal imagery that captures the phenological changes of the different land cover categories (Oetter, Cohen,

* Corresponding author. Tel.: +49 30 2093 6877; fax: +49 30 2093 6848.
E-mail address: cornelius.senf@geo.hu-berlin.de (C. Senf).

http://dx.doi.org/10.1016/j.rse.2014.10.018
0034-4257/© 2014 Elsevier Inc. All rights reserved.

Berterretche, Maiersperger, & Kennedy, 2001; Prishchepov, Radeloff, Dubinin, & Alcantara, 2012; Zheng, Campbell, & de Beurs, 2012).

The Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper (ETM+) sensors have proven to be a reliable source of data for land cover mapping because they have a relatively high spatial resolution of 30 m, a wide swath (185 km), and a frequent repeat cycle of 16 days (Gottschalk, Huettmann, & Ehlers, 2005; Wulder et al., 2008). Moreover, with Landsat 4 (1982–1993), Landsat 5 (1984 to 2013), Landsat 7 (since 1999), and Landsat 8 (since 2013), we now have more than 30 years of systematic Earth observations at a spatial resolution of 30 m, making Landsat a valuable source of data for long-term monitoring studies (Wulder, Masek, Cohen, Loveland, & Woodcock, 2012). Combining data from different Landsat sensors increases Landsat's effective temporal resolution to potentially 8 days (e.g. Landsat 5 and 7). However, most regions have not been continuously imaged every 16 days. Data acquisition from Landsat 7 follows the Long-term Data Acquisition Plan (LTAP), which schedules acquisitions based on seasonality, solar zenith angle, and cloud cover, among others, and gives priority to scenes over the continental U.S. (Arvidson, Goward, Gasch, & Williams, 2006). Further considering cloud cover and Landsat 7's failed scan-line corrector (SLC-off; since 31/05/2003), the temporal resolution is often much greater than 16 days. In practice, restrictions in data availability are likely when two or more images are needed, especially if an annual monitoring is foreseen (Ju & Roy, 2008).

Besides data availability, data accessibility also comes into play. In 2008, the United States Geological Survey (USGS) made all archived and new Landsat data freely available through the Earth Resources Observation and Science (EROS) Center (Wulder et al., 2012). The European Space Agency (ESA), being the European International Cooperator, holds approximately two million images that cover Europe and North Africa. To adapt the "Joint Principles for a GMES Sentinel Data Policy" (European Space Agency, 2009), the ESA revised its data policy in 2010, following the USGS and making all Landsat images available free of charge. However, access options still lag behind the EROS, and requirements include a brief project description that must be approved by ESA. After approval, data are processed by ESA and delivered to the user. To ease access, ESA will gradually process all data into an online archive for improved and faster access. At the time this manuscript was written, approximately 150,000 images acquired by the Kiruna ground station in northern Sweden were already available for direct download.

Some studies mapping land cover for habitat monitoring made use of sensors such as the MODerate-resolution Imaging Spectroradiometer (MODIS), the Advanced Very High Resolution Radiometer (AVHRR), or the Satellite Pour l'Observation de la Terre VEGETATION (SPOT VEGETATION), all of which have a coarser spatial resolution (250 to 1000 m), a wider swath (>2000 km), and more frequent repeat cycles (up to daily). However, the coarser spatial resolution limits their capability for monitoring habitats in spatially complex landscapes (Kerr & Ostrovsky, 2003; Osborne, Alonso, & Bryant, 2001; Wang, Franklin, Guo, He, & McDermid, 2009).

The Spatial and Temporal Adaptive Reflectance Fusion Model (STARFM) by Gao, Masek, Schwaller, and Hall (2006) has been proposed to make use of the strengths of different sensors — i.e. the 30 m spatial resolution of Landsat and the up-to-daily temporal resolution of MODIS. STARFM was previously used to improve single-date Landsat classifications of conservation tillage (Watts, Powell, Lawrence, & Hilker, 2011). In the study of Watts et al. (2011), adding STARFM-simulated imagery to single-date Landsat-based classifications increased the overall accuracy from 87.2% to 93%. However, it is yet unclear if STARFM will also help to improve classifications of spatio-temporal complex landscapes such as pseudo-steppes and how the results would compare to using multi-seasonal Landsat data.

The objectives of this study were to test whether and by what degree land cover maps of complex pseudo-steppe landscapes improve by integrating multi-seasonal Landsat imagery, as well as whether STARFM-simulated imagery can be used whenever original multi-seasonal Landsat observations are unavailable. To accomplish these objectives we developed five classification scenarios: 1) Using solely single-date Landsat imagery; 2) using single-date Landsat and STARFM-simulated imagery at key phenological dates; 3) using single-date Landsat imagery and a full STARFM-simulated time series; 4) using multi-seasonal Landsat data at key phenological dates; and 5) using all available Landsat data. Key phenological dates were identified using phenological profiles from MODIS Normalized Difference Vegetation Index (NDVI) time series. The multi-seasonal Landsat scenarios (Scenarios 4 and 5) provided a baseline for benchmarking the STARFM-based models. Moreover, we aimed to evaluate how temporal sampling of Landsat data influences the classification accuracies.

2. Study site and data

2.1. Study site

The study site is located in the Beja district in southern Portugal (Fig. 1), and covers approximately 2000 km^2 around the town of Castro Verde. The climate is Mediterranean, with hot summers (30–35 °C on average in July) and moderate cold winters (5–8 °C on average in January). Approximately 75% of the annual rainfall (500–600 mm) occurs between October and March. The pseudo-steppes of Castro Verde hold nationally and internationally important populations of endangered steppe bird species such as the great bustard (*Otis tarda*), little bustard (*Tetrax tetrax*), and black-bellied Sandgrouse (*Pterocles orientalis*; (Moreira et al., 2007). These populations led to the creation of a Special Protection Area for birds in 1999 (SPA), which covers an area of approximately 80,000 ha (Fig. 1).

The regional steppe-bird community habitat guilds were detailed by previous studies and correlated to specific land cover categories (Leitão et al., 2010; Moreira et al., 2007). We consequently focus on the following five land cover categories as defined by the previous studies: (1) Bare soils (i.e. plowed fields); (2) cereal fields; (3) grasslands; (4) shrublands; and (5) woodlands. Cereal fields, grasslands and bare soils are of special importance because they present the core classes of protection under the SPA. All other land cover categories (i.e. water and settlements) were excluded from the analysis using a mask from previous studies.

2.2. Landsat data

The average Landsat data availability over Castro Verde (WRS-2 path/row: 203/34) is three images per year from 2001 to 2012 (Table A1 in Appendix A). However, for many years there is only one scene or no scenes with cloud cover less than 10% and without SLC-off available in the USGS archive (2004–2006, 2008 and 2012), whereas in other years clouds present an over-average availability (e.g. six images in 2011). In the ESA archive, only images without geo-location and terrain-correction are available for our study area (Table A1 in Appendix A).

We acquired six available cloud-free L1T Landsat 5 TM images for the year 2011 from the USGS Landsat archive (Fig. 2). Using the high image availability in 2011 as the experimental set-up, we tested different scenarios of Landsat data availability by varying the Landsat dates used for calibrating the STARFM model, and also those used as input for the land cover classifications. This allows us to draw conclusions about the suitability of the method for mapping land cover in years where only few images without SLC-off and low cloud cover are available (e.g. 2004–2005/2008; Table A1 in the Appendix A). All Landsat images were atmospherically corrected to surface reflectance using the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) algorithm (Masek et al., 2006). We used all spectral bands, except for band 1 (blue band) and the thermal band (band 6) because of their lower sensitivity to vegetation properties and the high correlation of band 1 with band 2 and band 3 (Kuemmerle, Damm, & Hostert, 2008).