Remote Sensing of Environment 218 (2018) 119–131

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Invasive tree species detection in the Eastern Arc Mountains biodiversity hotspot using one class classification

Rami Piiroinen[a,b,*], Fabian Ewald Fassnacht[c], Janne Heiskanen[a,b], Eduardo Maeda[a,d], Benjamin Mack[e], Petri Pellikka[a,b,f]

[a] Earth Change Observation Laboratory, Department of Geosciences and Geography, University of Helsinki, P.O. Box 64, FI-00014 Helsinki, Finland
[b] Institute for Atmospheric and Earth System Research, Faculty of Science, University of Helsinki
[c] Institute of Geography and Geoecology, Karlsruhe Institute of Technology, Kaiserstraße 12, 76131 Karlsruhe, Germany
[d] Ecosystems and Environment Research Programme, Faculty of Biological and Environmental Sciences, University of Helsinki, P.O. Box 68, FI-00014 Helsinki, Finland
[e] GAF AG, Arnulfstr. 199, 80634 Munich, Germany
[f] Helsinki Institute of Sustainability Science, University of Helsinki

ARTICLE INFO

Keywords:
Biased support vector machine
Invasive tree species
Imaging spectroscopy
Africa
One class classification

ABSTRACT

Eucalyptus spp. and *Acacia mearnsii* are common exotic tree species in eastern Africa that have shown (strong) invasive behavior in some regions. *Acacia mearnsii* is considered a highly invasive species that is replacing native species and *Eucalyptus* spp. are known to consume high amounts of groundwater with suspected effects on native flora. Mapping the occurrence of these species in the Taita Hills, Kenya (part of the Eastern Arc Mountains Biodiversity Hotspot) is important as there is lack of knowledge on their occurrence and ecological impact in the area. Mapping methods that require a lot of fieldwork are impractical in areas like the Taita Hills, where the terrain is rugged and the infrastructure is poor. Our aim was hence to map the occurrence of these tree species in a 100 km^2 area using airborne imaging spectroscopy and laser scanning. We used a one class biased support vector machine (BSVM) classifier as it needs labeled training data only for the positive classes (*A. mearnsii* and *Eucalyptus* spp.), which potentially reduces the amount of required fieldwork. We also introduce a new approach for parameterizing and setting the threshold level simultaneously for the BSVM classifier. The second aim was to link the occurrence of these species to selected environmental variables. The results showed that the BSVM classifier is suitable for mapping *Acacia mearnsii* and *Eucalyptus* spp., holding the potential to improve the efficiency of field data collection. The introduced parametrization/threshold selection method performed better than other commonly used approaches. The crown level F1-score was 0.76 for *Eucalyptus* spp. and 0.78 for *A. mearnsii*. We show that *Eucalyptus* spp. and *A. mearnsii* trees cover 0.8% and 1.6% of the study area, respectively. Both species are particularly located on steeper slopes and higher altitudes. Both species have significant occurrences in areas close to the biggest remaining native forest patch (Ngangao) in the study area. Nonetheless, follow-up studies are needed to evaluate their impact on the native flora and fauna, as well as their impact on the water resources. The maps created in this study in combination with such follow-up studies could serve as base data to generate guidelines that authorities can use to take action in handling the problems these species are causing.

1. Introduction

A plant species is considered non-native or exotic if it is found in an ecosystem where it did not evolve. On the other hand, invasive plant species are defined as non-native plants that produce reproductive offspring in large numbers and at considerable distances from parent plants (Richardson et al., 2000). Woody plants, in general, were not widely recognized as invasive species until fairly recently (Richardson and Rejmánek, 2011). In contrast, nowadays invasive trees and shrubs are considered in some cases among the most conspicuous and damaging life-forms, threatening local flora and fauna. In Africa, some alien tree species serving as backbone of the local plantation forestry have high economic significance, but at the same time decimate land and water resources (Chenje and Mohamed-Katerere, 2006).

* Corresponding author at: Earth Change Observation Laboratory, Department of Geosciences and Geography, University of Helsinki, P.O. Box 64, FI-00014, Helsinki, Finland.
E-mail addresses: rami.piiroinen@helsinki.fi (R. Piiroinen), fabian.fassnacht@kit.edu (F.E. Fassnacht), benjamin.mack@gaf.de (B. Mack).

https://doi.org/10.1016/j.rse.2018.09.018
Received 22 May 2018; Received in revised form 18 September 2018; Accepted 21 September 2018

One of the most invasive alien tree species in Africa is *Acacia mearnsii*, featured also in the list of '100 of the World's Worst Invaders' (Lowe et al., 2000). This species, native in Australia, has been shown to compete with native species, to reduce native biodiversity, and to reduce water availability in riparian zones (Boudiaf et al., 2013; Richardson and Rejmánek, 2011). For instance, in South-Africa *A. mearnsii* was originally planted on 107,000 ha, but is estimated now to have spread to a total area of 2,500,000 ha (Nyoka, 2003). Another genus of trees known to cause environmental problems in sub-Saharan Africa and also considered invasive in some areas is *Eucalyptus* (Richardson and Rejmánek, 2011). For instance, conversion of grassland by afforestation with alien *Eucalyptus* spp. affects negatively the catchment runoff (Turpie et al., 2008). In some cases *Eucalyptus* spp. plantations have even completely dried up rivers (Rodriguez-Suarez et al., 2011). The leaf litter of *Eucalyptus* spp. (including *Eucalyptus saligna*) also contain phytotoxic compounds, that inhibit germination and initial growth of certain grassland species, and possible allelopathic effects (Silva et al., 2017).

While the adverse impacts of *Eucalyptus* spp. and *A. mearnsii* are well understood in South Africa (Nyoka, 2003; Turpie et al., 2008), fewer assessments have been conducted elsewhere in Africa. For instance, in Kenya, detailed maps of the current occurrences of these invasive species are still missing. In this study, we address this research gap by developing an efficient approach for assessing the occurrence of *A. mearnsii* and *Eucalyptus* spp. in the Taita Hills, Kenya, with limited field data.

The Taita Hills are part of the Eastern Arc Mountains biodiversity hotspot, which is known to host many endemic species (Burgess et al., 2007). However, according to a recent study, only 0.8% of the Taita Hills region are still covered with the native indigenous cloud forests which contain a large share of the endemic species and biodiversity of the area (Thijs et al., 2015). Many exotic tree species have been introduced to produce lumber (*Eucalyptus* spp., *Grevillea robusta*), tannin (*A. mearnsii*) and food (*Mangifera indica*, *Persea americana*). Aside from pure plantation forests, tree cover has increased on the croplands as treeless fields have been converted to agroforestry systems (Pellikka et al., 2018). These agroforestry systems often include exotic tree species with some of them being considered invasive. Mapping the occurrence of these invasive species would hence be highly valuable as the current spread and the impact of these species on the ecosystem in the study area is not well known. One existing study on the occurrence of tree species in the Taita Hills was based on field sampling (Thijs et al., 2015), which is an accurate but time-consuming approach that is not a practical solution for mapping the species at a broad scale and with high spatial accuracy.

An alternative approach for inventorying tree species over the entire region is provided by remote sensing (RS) techniques (Fassnacht et al., 2016). Imaging spectroscopy (IS) and airborne laser scanning (ALS) are the most common RS data sources used for the classification of tree species in the research literature (Fassnacht et al., 2016). Using these two data sources together (data fusion) has yielded the best results in many cases (Fassnacht et al., 2016). The data fusion is often performed at object level as it enables calculating smooth spectral features but also features that depict the structure and shape of a tree. However, in the tropics, the canopy structure is often complex and the automatic delineation of tree crowns is challenging (Feret and Asner, 2013; Piiroinen et al., 2017). Thus, pixel-based mapping approaches have also been presented (Baldeck et al., 2015). Most of the studies utilizing IS and ALS data for mapping tree species have been conducted in temperate forests, while fewer studies have been located in tropical or sub-tropical areas, and those mainly in Central America, South America and southern Africa (Baldeck and Asner, 2015; Baldeck et al., 2015; Cho et al., 2012; Fassnacht et al., 2016; Graves et al., 2016). Only one recent study was conducted in Kenya (Piiroinen et al., 2017).

Tree species classification and mapping studies typically use supervised classification approaches, where the classifier is trained using field measurements of all the tree species that are present in the study area. This approach is sometimes impractical. This particularly applies for tropical regions where a single study site may have dozens or hundreds of different tree species, which makes collecting representative training and validation data very laborious, particularly in areas with limited infrastructure. Furthermore, only a few species might be relevant for the application or research question. If the latter applies, the use of a one class classification (OCC) approach, where labeled data is needed only for the positive class (that is, a single tree species) might be an efficient alternative (Mũnoz-Marí et al., 2010).

In RS studies, OCCs have been used, for example, to detect focal tree species in tropical rainforests (Baldeck et al., 2015), Natura 2000 habitats and high nature value grassland habitats (Stenzel et al., 2014, 2017), for invasive species detection (Skowronek et al., 2017a, 2017b), and detecting savanna tree species in Africa (Baldeck and Asner, 2015). From the plethora of available OCC algorithms, particularly one class support vector machine (OCSVM), biased support vector machine (BSVM) and Maxent have been frequently used (Mack and Waske, 2017). OCSVM (Scholkopf et al., 1999) uses only data from the class of interest to train the classifier, while BSVM is a semi-supervised classification algorithm that utilizes also unlabeled samples (Liu et al., 2003). In a recent study conducted in Panama, very high classification accuracies were achieved with BSVM for detecting non-flowering focal tree species at the pixel level (Baldeck et al., 2015). Mack and Waske (2017) showed in their comparison of different OCC algorithms that BSVM had the highest discriminative potential followed by Maxent (with parameter tuning), Maxent (with default parameters) and OCSVM. Similarly, Stenzel et al. (2017) reported that BSVM outperformed Maxent (with default parameters) and OCSVM in the classification of high nature value grassland areas. Stenzel et al. (2017) concluded that the results could have been further improved by more sophisticated parameter tuning.

One of the benefits of using OCCs in invasive species mapping is that many governmental organizations in charge of nature conservation and management keep record of the known locations of certain invasive species and this information can be readily used to make initial maps. For instance, Wakie et al. (2014) collected 143 observations of invasive *Prosopis juliflora* in Ethiopia using targeted field sampling that was based on the pre-existing knowledge of heavily infested sites that the local communities and government employees had. Wakie et al. (2014) then used MODIS data and Maxent to model the occurrence of *P. juliflora*. The same approach does not work in supervised classification methods as information on all other species (the negative class) is often missing.

The first aim of this study was to examine the potential of the OCC approach and a BSVM classifier in combination with pixel-level data fusion of IS and ALS data for mapping common invasive tree species, namely *Eucalyptus* spp. and *A. mearnsii*, in the Taita Hills, Kenya. We selected BSVM (Liu et al., 2003) based on its good performance in previous studies (Mack and Waske, 2017; Stenzel et al., 2017). Furthermore, Mack et al. (2014) suggested that better results can be achieved for BSVM when the threshold (cut-off value) is selected manually based on diagnostic plots in case the automatic procedures fail. However, in practice, the manual tuning of the threshold might be challenging. To respond to this challenge, we introduce a new approach for selecting the model and tuning the threshold simultaneously.

The second aim was to map the occurrence of these species and relate their occurrence with selected environmental variables. The goal was to achieve a better understanding of the locations that are most heavily affected by these species. The results can serve as a baseline for studying the impacts of these species on the ecosystem, biodiversity and water resources.