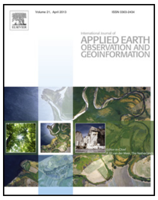
International Journal of Applied Earth Observation and Geoinformation 28 (2014) 90–101

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

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



Using multiple Landsat scenes in an ensemble classifier reduces classification error in a stable nearshore environment

Anders Knudby[a,*], Lina Mtwana Nordlund[b], Gustav Palmqvist[c], Karolina Wikström[c], Alan Koliji[c], Regina Lindborg[d], Martin Gullström[c]

[a] Department of Geography, Simon Fraser University, 8888 University Drive, BC V5A1S6, Canada
[b] Western Indian Ocean - Community, Awareness, Research, and Environment (WIO CARE), P.O. Box 4199, Zanzibar, Tanzania
[c] Department of Ecology, Environment and Plant Sciences, Stockholm University, S-106 91 Stockholm, Sweden
[d] Department of Physical Geography and Quaternary Geology, Stockholm University, SE-106 91 Stockholm, Sweden

ARTICLE INFO

Article history:
Received 21 August 2013
Accepted 22 November 2013

Keywords:
Remote sensing
Classification
Landsat
Ensemble classifier
Random forest
Nearshore

ABSTRACT

Medium-scale land cover maps are traditionally created on the basis of a single cloud-free satellite scene, leaving information present in other scenes unused. Using 1309 field observations and 20 cloud- and error-affected Landsat scenes covering Zanzibar Island, this study demonstrates that the use of multiple scenes can both allow complete coverage of the study area in the absence of cloud-free scenes and obtain substantially improved classification accuracy. Automated processing of individual scenes includes derivation of spectral features for use in classification, identification of clouds, shadows and the land/water boundary, and random forest-based land cover classification. An ensemble classifier is then created from the single-scene classifications by voting. The accuracy achieved by the ensemble classifier is 70.4%, compared to an average of 62.9% for the individual scenes, and the ensemble classifier achieves complete coverage of the study area while the maximum coverage for a single scene is 1209 of the 1309 field sites. Given the free availability of Landsat data, these results should encourage increased use of multiple scenes in land cover classification and reduced reliance on the traditional single-scene methodology.

1. Introduction

Land cover mapping is one of the most common applications of remote sensing. In its supervised form, field observations are used to train a classifier to predict the land cover of an area from its spectral radiance or reflectance, texture (Purkis et al., 2006) and, in object-based classification, the shape, size and context of image segments (Blaschke, 2010). These values are typically derived from a single scene of remote sensing data, or from a combination of scenes acquired from different sensors (data fusion, e.g. Bejarano et al., 2010). For tropical nearshore areas, the focus of this study, supervised classification has been used extensively with passive optical satellite and airborne data to map both terrestrial and submerged areas, including ecologically and economically important mangrove forests, seagrass beds, and coral reefs (Green et al., 2000). Submerged but optically shallow areas have typically been defined on the basis of either reef geomorphology (Purkis et al., 2010; Andréfouët and Guzman, 2005; Smith et al., 1975; Suzuki et al., 2001) or dominant substrate (Green et al., 1996; Ahmad and Neil, 1994; Purkis et al., 2002), and the relationship between sensor spatial and spectral properties, the number of classes mapped, and the achievable map accuracy has been explored for commonly used satellite sensors as well as airborne hyperspectral instruments (Mumby et al., 1997). Despite difficulties in comparing results across the relevant spatial resolutions (Capolsini et al., 2003), collectively results suggest that availability of one or more bands operating in the 400–500 nm wavelength range improves accuracy (Hedley et al., 2012), high spatial resolution improves accuracy (Andréfouët et al., 2003; Mumby and Edwards, 2002), high spectral resolution improves accuracy when relatively few classes are mapped (Capolsini et al., 2003), and airborne hyperspectral sensors, which combine very high spatial and spectral resolution, consistently outperform satellite sensors (Mumby et al., 1997; Kutser et al., 2003; Botha et al., 2013). Many studies have adopted the maximum likelihood algorithm as a standard supervised per-pixel classifier, and rigorous comparisons of a wider range of per-pixel classifiers, as has been carried out elsewhere (Pal and Mather, 2005; Pal, 2005; Brenning, 2009; Duro et al., 2012), has not been performed for tropical nearshore areas. However, it has been demonstrated that the use of contextual editing can improve classification accuracy by providing additional information necessary to separate spectrally similar substrate types, either

---

[*] Corresponding author. Tel.: +1 778 782 3876.
E-mail address: aknudby@sfu.ca (A. Knudby).

0303-2434/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jag.2013.11.015

by re-classing areas in an initial classification result (Benfield et al., 2007), or through pre-segmentation (e.g. by geomorphologic zone) and independent classification of individual segments (Andréfouët et al., 2003; Suzuki et al., 2001). Several studies have also compared the obtainable accuracies between a per-pixel classifier and Object-Based Image Analysis (OBIA) approaches and concluded, despite problems associated with direct comparisons, that OBIA produces substantially more accurate map products (Phinn et al., 2012; Benfield et al., 2007; Leon and Woodroffe, 2011).

High-resolution multispectral satellite sensors (IKONOS, QuickBird, WorldView-2) combine the presence of one or more bands in the 400–500 nm wavelength range with high spatial resolution while being more affordable than airborne hyperspectral data, and these sensors have therefore increasingly been used for mapping of small tropical nearshore areas. However, with a few exceptions (Purkis et al., 2010; Rowlands et al., 2012) Landsat TM/ETM+ data have remained the mainstay of large-scale mapping efforts (UNEP/WCMC, 2010; Andréfouët et al., 2006) due to the large extent of their scenes and their free availability. For the same reasons, the Landsat archive is still the best source of remote sensing data in parts of the world where at-cost satellite data are not a feasible option for mapping projects. However, use of Landsat data for mapping large tropical nearshore areas is challenging for a number of reasons. Data availability is limited because Landsat 5 TM has not produced data for large parts of the world since 2011, Landsat 7 ETM+'s scan line corrector (SLC) malfunctioned on May 31, 2003, after which approximately 22% of the data from each ETM+ scene has been missing, and Landsat 8 has only been producing data since April 11, 2013. Although some continuous coverage has thus been provided, Landsat 7 ETM+ with its SLC malfunction has been the only source of Landsat data for a significant period of time. The frequent presence of clouds in the tropics further limits the availability of useful Landsat scenes, and identification of a single appropriate Landsat scene for a large area of interest may therefore be difficult or impossible. For example, not a single Landsat TM/ETM+ scene exists for the area used in this study (Zanzibar Island, Path 166 Row 64), in which the entire nearshore area is free of clouds. This situation is not unique to Zanzibar, and is problematic because the methodology used in the large majority of studies mapping tropical nearshore environments is based on the presence of a single cloud-free scene covering the entire study area. This methodology can be appropriate for small-scale studies for which cloud-free scenes can be obtained (Andréfouët, 2008), but fails with increasing frequency as the size of the study area increases. For remote sensing to realize its full potential in mapping tropical nearshore areas, an alternative methodology that allows utilization of cloud- and error-affected scenes is therefore needed.

For regional- and global-scale terrestrial land cover mapping with coarse-resolution satellite data, the presence of clouds is dealt with by combining cloud- and shadow-free pixels from several scenes to create a cloud-free composite, which then forms the basis for land cover classification (Eva et al., 2004; Latifovic et al., 2004; Bartalev et al., 2003). The success of compositing is heavily dependent on the identification of an appropriate compositing algorithm (Cihlar et al., 1994) used to identify the best pixel from two or more candidate scenes, however, no effective compositing criteria exist for nearshore environments so an alternative is necessary. Using two Landsat scenes, Leon and Woodroffe (2011) modified the compositing approach by identifying clouds and cloud shadows in a master scene and then replaced the affected pixels with those from a second scene to create a cloud-free composite scene. In addition to accurate cloud- and shadow-detection, the success of this approach depends on effective radiometric normalization, which is likely to be difficult for nearshore areas because these typically comprise a small part of the scene for which normalization statistics are derived and applied, and because differences in tidal stage, turbidity and sea state between scene acquisitions can produce large effects on the observed reflectance spectrum. The presence of residual radiometric effects inherited from individual scenes will reduce accuracy when a single classification algorithm is applied to the composite scene, although the use of OBIA approaches may be effective in reducing this effect (Zhou et al., 2009). More importantly, the use of a single cloud-free composite does not utilize the information available in the unused but cloud-free pixels. In this study we develop and test an alternative approach: the use of an ensemble classifier that relies on classification results from multiple cloud- and error-contaminated Landsat scenes, all acquired within a period of time in which relevant land cover change can be assumed negligible. We compare a range of statistical and machine-learning classification algorithms, assess the number of scenes needed to optimize the overall accuracy classification accuracy, and assess the influence of cloud shadow detection, balancing of the training data, and use of distance-based contextual features on classification accuracy. We demonstrate how the use of multiple Landsat scenes in an ensemble classifier can achieve both a complete coverage of the study area and substantially improve classification accuracy compared to the traditional single-scene approach.

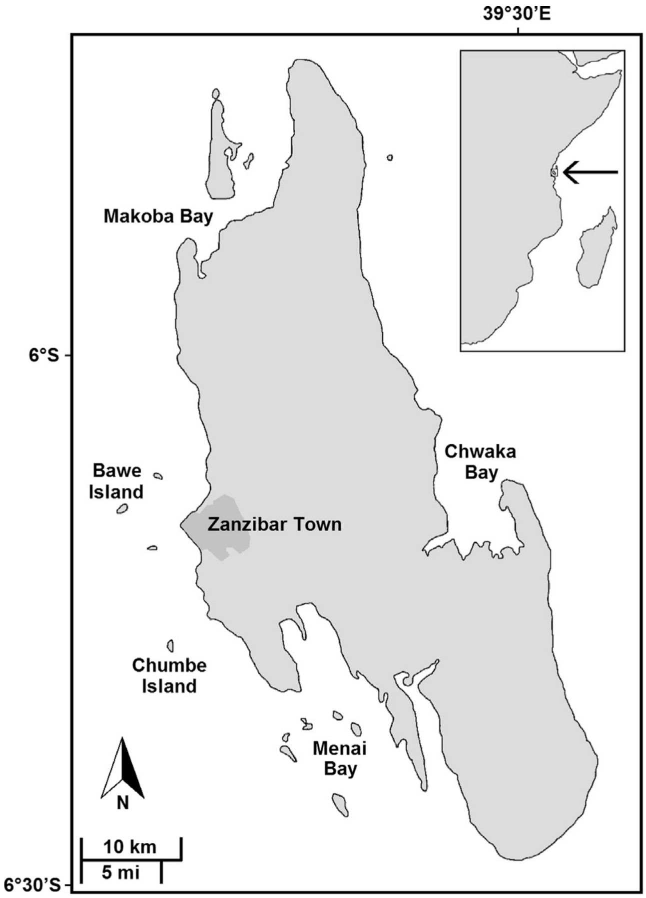


Fig. 1. Zanzibar Island, its three main bays, and Chumbe and Bawe islands. Insert shows Zanzibar Island's location off the East African coast.

2. Study area

Zanzibar Island (Unguja) is located in the Western Indian Ocean, 30 km from the Tanzanian mainland, and is surrounded by numerous islets (Fig. 1). Its climate is tropical with two main seasons characterized by northern winds and high temperatures from November to March and southern winds and cooler temperatures