Available online at www.sciencedirect.com

SCIENCE DIRECT®

Remote Sensing of Environment 95 (2005) 428–443

Remote Sensing of Environment

www.elsevier.com/locate/rse

A new GLM-based method for mapping tree cover continuous fields using regional MODIS reflectance data

Markus Schwarz, Niklaus E. Zimmermann*

Swiss Federal Research Institute WSL, Dept. of Landscape Research, Zuercherstrasse 111, 8903 Birmensdorf, Switzerland

Received 27 April 2004; received in revised form 5 November 2004; accepted 16 December 2004

Abstract

Knowledge about land cover and its change is an important input for the monitoring and modeling of ecological and environmental processes from the regional to the global scale. Considerable efforts have been made to develop global continuous fields for different land cover types at large spatial scales based on NOAA-AVHRR and TERRA-MODIS data and a range of techniques have been applied to depict the sub-pixel fraction of land cover types from these data. In this study, a new methodology is described for deriving and optimizing continuous fields of tree cover for complex topography at the regional scale of the European Alps using generalized linear models (GLM). MODIS data (MOD09) at a spatial resolution of 500 m were used to calibrate the models against regional training data of fractional tree cover. For evaluating the method we test the GLM model output to a regression tree model (using the same data structure). Further we test the resulting GLM-based tree cover continuous fields against two different, independent test data sets; one of which is spatially separated and the other is from within the calibration area. Finally, we compare the GLM model output with two available global data sets at spatial resolutions of 1 km and 3 km: (1) TERRA-MODIS Vegetation Continuous Fields product (MOD44), and (2) the NOAA-AVHRR vegetation continuous fields. Our GLM-based method results in high accuracy (MAE=9.1%) and low bias (−1.2%) across the combined evaluation and calibration area, and with small differences only between the calibration and the spatially separated evaluation area (1.3%). Compared to the regression tree model the results from the GLM model for all analyses are significantly better. Thus we conclude that generalized linear models are appropriate for deriving continuous fields of fractional tree cover for complex topography at the regional scale. GLMs can handle nonlinear relationships present in the training data set well, and the method is robust with respect to sample size and the number of months used for calibration. Regional calibrations of vegetation continuous fields may offer significantly improved predictions compared to globally calibrated models. Such regionally calibrated and optimized models may serve as valuable tools for regional monitoring of land cover pattern and its temporal change.

Keywords: European Alps; Fractional cover; Generalized linear models; GLM; Multi-temporal; MODIS; Tree cover continuous fields; Weighted kappa

1. Introduction

The environment of the European Alps with its characteristic complex topography and its fine-scale and often traditional land use regime is considerably exposed to both natural environmental threats and human impacts and exploitation (Tasser & Tappeiner, 2002). The combined pressure on the alpine environment arising from past development, rapid land use change, tourism, and possible climate change (Barry, 1994; Grabherr et al., 1994; Theurillat & Guisan, 2001) is high when compared to other similar environments (Price & Haslett, 1995; Tasser & Tappeiner, 2002). According to the 1989 International Convention on the Protection of the Alps, mountain regeneration and sustainability are key issues. For sustainable use and management, and for large-area monitoring and modeling of these Alpine terrestrial habitats, consistent land cover information across national borders is indispensable (CIPRA, 2001). Additionally, such information is required for many aspects of earth sciences and analyses of global

---

* Corresponding author.
E-mail address: niklaus.zimmermann@wsl.ch (N.E. Zimmermann).

0034-4257/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2004.12.010

change, including resource management or biodiversity assessment (Townshend et al., 1994). The knowledge of spatial patterns and extent of terrestrial ecosystems is further important for the evaluation and management of carbon sinks and sources at national and continental levels, as land vegetation stores considerable amounts of comparably mobile carbon stocks (Cernusca et al., 1998). Tree cover mapping has particularly grown in importance as a result of the need to quantify the global woody biomass (Hansen et al., 2002). Finally, the knowledge of the distribution of tree cover is an important input to modeling of biogeochemical cycles and climate feedbacks (Sellers et al., 1997; Townshend et al., 1994). Thus, reliable land cover information and data sets are needed for a wide range of environmental monitoring applications. Previous studies have shown that remote sensing based data are appropriate tools to detect land cover change (Justice et al., 1998; Rogan et al., 2002).

Considerable efforts have recently resulted in the development of global and continental land cover data at small spatial scales based on the Advanced Very High Resolution Radiometer (NOAA-AVHRR) and the Moderate Resolution Imaging Spectroradiometer (TERRA-MODIS). Examples include: (1) the PELCOM land cover map (Mucher et al., 2000), (2) the UMD global land cover map (Hansen & Reed, 2000), (3) the IGBP-DIS continental and global land cover map (Belward et al., 1999; Loveland & Belward, 1997; Townshend et al., 1994), and (4) the MODIS MOD12 global land cover product (Friedl et al., 2002). The paradigm for describing the characteristics of the surface covered by these data sets is to classify each pixel into a land cover type based on a predefined classification scheme (DeFries & Townshend, 1994; DeFries et al., 1998). However, this approach has certain limitations (Aplin & Atkinson, 2001). In most cases, all degrees of mixing of pure land cover classes within a pixel can be found due to the continuum of variation found in the landscape (Foody & Hill, 1996) and the mixed nature of land cover at coarse (≈1 km) spatial resolution (Schowengerdt, 1996). This is especially true for complex landscapes of mountainous terrain and for small-scale structured landscape of traditional land use schemes with its typically small patches of human and natural disturbances. Hence, the discretization of land cover into a limited number of categories results in a loss of information (Ju et al., 2003). This loss of information can have significant impact on subsequent modeling as has been demonstrated, e.g., by Pierce and Running (1995). Alternatively, a number of techniques to depict the subpixel fraction of landscape components from the same remotely sensed data have been applied. In general, such techniques make use of high temporal resolution to overcome the restrictions of limited spatial resolution. Statistical approaches used so far (Fernandes et al., 2004; Hansen et al., 2002) are: (1) fuzzy membership functions (Foody, 1994; Foody & Cox, 1994), (2) artificial neural networks (Atkinson et al., 1997; Braswell et al., 2003), (3) regression trees (DeFries et al., 1997), (4) decision trees (McIver & Friedl, 2002), and (5) linear mixture models (Adams et al., 1995). The resulting continuous field maps–containing the fraction of landscape components as a continuous variable–offer the advantage of summarizing the effects of spatial heterogeneity better than the discrete land cover maps based on the same data sources. As a result, such products have a higher potential to accurately monitor land cover change over time (Hansen et al., 2002).

Several attempts have recently been made to map fractional landscape components using moderate resolution imagery covering large areas. First investigations using AVHRR data made use of a linear mixture model to derive global continuous fields of multiple vegetation classes (DeFries et al., 1999). Recent efforts have shown that linear mixture models may not be suitable in cases when multiple scattering results in nonlinear mixing (Ju et al., 2003). In this context nonlinear decision rules may produce better results. Hansen et al. (2002) presented an improved technique to derive the fraction of tree cover per pixel using a combination of a regression tree algorithm and linear least-square models. The results showed that this method is an improvement over the linear-mixture model and it can better handle the nonlinear relationships present in a global sample of tree cover.

The goal of this paper is to present a new promising approach for fractional cover mapping of moderate resolution imagery based on generalized linear models (GLM) and using regional training data. GLMs are an extension of the linear (least-square regression) modeling that allows models to be fitted to data with errors following other than (only) Normal distributions, and for dependent variables following other than a Normal distribution, such as the Poisson, Binomial and Multinomial (McCullagh & Nelder, 1989). Ordinary least-squares regression assumes, e.g., that the model response varies continuously and that it is unbounded. GLMs of the binomial model family overcome this difficulty by linking the binary response to the explanatory covariates through the probability of either outcome, which varies continuously from 0 to 1 (Dobson, 2002). This approach is often referred to as logit regression. Other model families allow fitting response variables of different restricting characteristics (Poisson regression, etc.). Tree cover fractions have the same restrictions as binary dependent variables (upper and lower bounds); except they vary continuously between 0 and 1. Thus, GLMs of the binomial model family may be equally well applied to fractional cover data as can be applied to presence/absence data. While the GLM output of a binary response variable is interpreted as "probability to occur", one can take the GLM output of a cover response at the original scale, which is cover fraction. GLMs are commonly used in environmental research (see Guisan & Zimmermann, 2000, for a review; Zimmermann & Kienast, 1999, for an application), though we found few examples with remote sensing data. One successful