Environmental Modelling & Software 83 (2016) 245–254

Contents lists available at ScienceDirect

Environmental Modelling & Software

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

Importance of spatially distributed hydrologic variables for land use change modeling

Paul D. Wagner*, Björn Waske

Remote Sensing and Geoinformatics, Institute of Geographical Sciences, Freie Universität Berlin, D-12249 Berlin, Germany

ARTICLE INFO

Article history:
Received 17 February 2016
Received in revised form
4 May 2016
Accepted 4 June 2016

Keywords:
Land use modeling
Land-Use/Cover Change (LUCC)
Logistic regression
Hydrologic modeling
India

ABSTRACT

Land use changes have a pronounced impact on hydrology. Vice versa, hydrologic changes affect land use patterns. The objective of this study is to test whether hydrologic variables can explain land use change. We employ a set of spatially distributed hydrologic variables and compare it against a set of commonly used explanatory variables for land use change. The explanatory power of these variables is assessed by using a logistic regression approach to model the spatial distribution of land use changes in a meso-scale Indian catchment. When hydrologic variables are additionally included, the accuracies of the logistic regression models improve, which is indicated by a change in the relative operating characteristic statistic (ROC) by up to 11%. This is mostly due to the complementarity of the two datasets that is reflected in the use of 44% commonly used variables and 56% hydrologic variables in the best models for land use change.

1. Introduction

Anthropogenic land cover transformation and land-use activities, particularly due to urbanization, farmland expansion and agricultural intensification, are among the most important components of global environmental change (Lambin et al., 2001; Turner et al., 2007). Regional land use and land cover change affects water resources (Foley et al., 2005) and vice versa, land use and land cover change is affected by changes in water resources (Lambin et al., 1999). Consequently, land use is a key component of hydrologic models (e.g., MIKE SHE, Refsgaard and Storm, 1995; SWAT, Arnold et al., 1998) and a key research topic in hydrology (DeFries and Eshleman, 2004). However, in studies and models of land use change, the potential explanatory power of hydrologic patterns for land use change is rarely exploited.

Several land use change models use spatially distributed explanatory variables to predict the spatial distribution of land use changes (e.g., CLUE, Verburg and Overmars, 2009; GEOMOD, Pontius et al., 2001). These explanatory variables for land use change can broadly be classified in biophysical variables (e.g., topography, soil characteristics, climatic variables) and socio-economic variables (e.g., population density, distance to roads or cities; Veldkamp and Lambin, 2001). Established variables with proven explanatory power include topographic variables, soil characteristics, population variables, and distance variables (e.g., Baumann et al., 2011; Mas et al., 2004; Oñate-Valdivieso and Bosque Sendra, 2010; Schneider and Pontius, 2001; Verburg et al., 2002, 2004). Sometimes also hydrologic variables like precipitation, available water capacity (Sohl et al., 2007), evapotranspiration (Prishchepov et al., 2013), a yearly moisture index (Rutherford et al., 2007), or a wetness index (Kim et al., 2014) are used to explain land use change. However, the distance to water bodies or streams often serves as the only proxy variable for water availability (e.g., Huang and Cai, 2007; Verburg et al., 2004; Xu et al., 2013), possibly due to the ease of computation. One reason that could have led to the rare consideration of hydrologic variables might be the limited availability of such spatially distributed variables. Moreover, a considerable amount of preprocessing is required e.g., to derive soil moisture patterns from satellite data (Koyama et al., 2010), evapotranspiration estimates using SEBAL (Bastiaanssen et al., 1998), or interpolated precipitation patterns (Buytaert et al., 2006). In summary, hydrologic variables are not among the well-established and commonly used variables in land use change modeling, even though they may affect land use change. Particularly in data scarce regions or in areas where few explanatory variables are available as spatially explicit data (Veldkamp and Lambin, 2001), a combination of different datasets to optimize information content seems promising. Moreover, due to the

* Corresponding author.
E-mail address: paul.wagner@fu-berlin.de (P.D. Wagner).

http://dx.doi.org/10.1016/j.envsoft.2016.06.005
1364-8152/

interdependence of hydrology and land use, the inclusion of hydrologic information in land use change models could be valuable.

In this study, we compare a set of commonly used biophysical and socio-economic variables to a set of hydrologic variables with regard to their potential to explain land use change. The study is carried out in an Indian catchment that experiences seasonally limited rainfall, where hydrologic variables have a particular importance with regard to land use changes (Lambin et al., 1999). Moreover, data availability for the study area is limited, so that additional spatially distributed information is valuable. To evaluate the potential of including hydrologic variables in land use change models we focus on three main objectives. We test i) if spatially distributed hydrologic variables can explain patterns of land use change in the study area, and ii) if they provide complementary information as compared to commonly used biophysical and socio-economic variables. Moreover, we iii) analyze and evaluate the explanatory variables and the derived probability maps for land use changes.

2. Materials and methods

2.1. Study area

The catchment of the Mula and Mutha Rivers is situated in western India upstream of the city of Pune. Major parts of the meso-scale catchment (2036 km^2) are located in the Western Ghats (Fig. 1). Elevation ranges from 550 m in Pune up to 1300 m a.s.l. on the top ridges in the Western Ghats. The climate is tropical wet and dry with a dependence of rainfall on the summer monsoon (June to October). Annual rainfall amounts decrease from approximately 3500 mm a^{-1} in the western part of the catchment to 750 mm a^{-1} in the eastern part of the catchment (Gadgil, 2002; Gunnell, 1997). The majority of the study area (92.5%) is covered by a sandy clay loam while the rest (7.5%) is covered by clay (Food and Agriculture Organization of the United Nations (FAO), 2003). Land use is dominated by semi-natural vegetation (70.7%), with forests mainly on the higher elevations in the west, whereas shrubland and grassland cover lower elevations (Fig. 1). Cropland (13.5%) is mainly found in proximity to water sources and settlements and is dominated by small fields (<1 ha) with rainfed agriculture during the monsoon season and irrigation during the dry season (rice-wheat rotation, sugarcane, mixed cropland). Typically, two crops per year are grown. Urban area (10.1%) is predominantly found in the eastern part of the catchment where the city of Pune and its surrounding settlements are located (Fig. 1). Reservoirs and rivers account for 5.7% of the study area (Wagner et al., 2013).

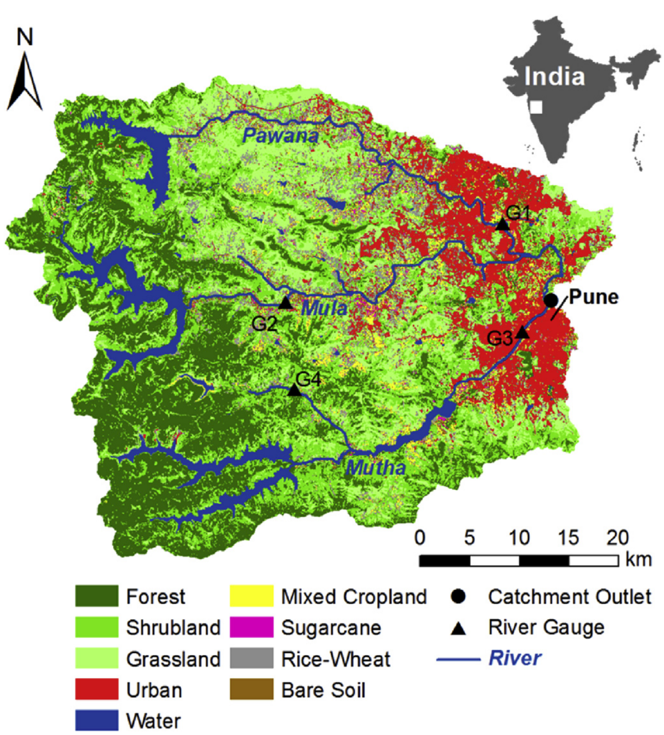

Fig. 1. Land use of the Mula and Mutha Rivers catchment in 2009/10.

2.2. Land use change

We use three land use maps for the cropping years 1989/90, 2000/01, and 2009/10 from Wagner et al. (2013) to determine land use changes between 1989/90 and 2000/01 and between 2000/01 and 2009/10, respectively. The maps are based on the classification of multispectral remote sensing data from different cropping seasons, using a stratified knowledge-based approach and a maximum likelihood classifier. The classification is challenging due to strong seasonal variations that result in two main growing seasons, various cropping practices and crop types, small-scale agriculture, and the limited availability of cloud-free satellite data. Therefore, the averaged overall accuracy per cropping year varies between 62% and up to more than 90%. While some classes are mapped with relative high accuracies (e.g., as sugarcane, rice, forest), classification accuracies for grassland, shrubland and mixed cropland are lower, compared to the accuracies achieved for other classes (Wagner et al., 2013). One reason for this might be that the differentiation between such land cover classes is often challenging, due to spectral ambiguity within the multispectral data (Song et al., 2002; Stefanski et al., 2014).

A post-classification change detection between the land use maps of 1989/90 and 2009/10 shows that the catchment has been affected by a pronounced urbanization from 5.1% to 10.1% of the study area as well as an expansion of farmland from 9.7% to 13.5%. Consequently, semi-natural vegetation has decreased from 79.8% to 70.7%. Due to lower class specific classification accuracies, changes within the semi-natural classes can partly be interpreted as pseudo change, e.g., 59.8% of changed shrubland areas and 79.5% of changed grassland areas in 2009/10 have been converted from the respective other class when compared to the land use of 1989/90 (Wagner et al., 2013).

2.3. Explanatory variables

Three sets of variables are tested to explain land use changes: i) commonly used biophysical and socio-economic variables in land use science, hereafter referred to as common variables, ii) modeled hydrologic variables, and iii) all variables together (variable sets i and ii). The first variable set consists of population density, distance to roads, distance to rivers, elevation, slope, aspect, and soil type (Table 1). These variables are widely used to explain land use changes (e.g., Oñate-Valdivieso and Bosque Sendra, 2010; Sohl et al., 2007; Verburg et al., 2002). We have used OpenStreetMap data from which the distance to roads variable has been calculated. The roads polygon shapefile (OpenStreetMap, 2015) captures the general characteristics of the road network quite well, even though some smaller roads are missing when compared to satellite data. Unfortunately, road networks for 1990 or 2000 have not been available to this study, so that changes in this variable are not represented. Moreover, we have employed a 30 m digital elevation model (DEM) based on ASTER satellite data which was processed and evaluated by Wagner et al. (2011). ASTER data has been proven