Forest Ecology and Management 310 (2013) 64–73

Contents lists available at ScienceDirect

Forest Ecology and Management

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

Combining ensemble modeling and remote sensing for mapping individual tree species at high spatial resolution

Robin Engler *, Lars T. Waser, Niklaus E. Zimmermann, Marcus Schaub, Savvas Berdos, Christian Ginzler, Achilleas Psomas

Swiss Federal Research Institute WSL, Zuercherstrasse 111, CH-8903 Birmensdorf, Switzerland

ARTICLE INFO

Article history:
Received 5 April 2013
Received in revised form 21 July 2013
Accepted 31 July 2013

Keywords:
Vegetation mapping
Aerial imagery
Species distribution modeling
Ensemble forecast
Forest ecosystems
Switzerland

ABSTRACT

The ability to map vegetation and in particular individual trees is a key component in forest management and long-term forest monitoring. Here we present a novel approach for mapping individual tree species based on ensemble modeling, i.e. combining the projections of several modeling techniques in order to reduce uncertainty. Using statistical modeling in conjunction with high-resolution aerial imagery (50 cm spatial resolution) and topo-climatic variables (5 m spatial resolution), we map the distributions of six major tree species (3 broadleaf and 3 conifers) in a study area of North-Eastern Switzerland. We also compare the relative predictive power of both topo-climatic and remote-sensing variables for mapping the spatial tree patterns and assess the importance of calibration data quality on model performance. We evaluate our projections using cross-validation as well as with independent data. Overall, the evaluations that we obtain for our vegetation maps are in line with, or higher than, those in similar studies. Depending on the considered tree species, 47.8–85.6% of our samples were correctly predicted, and we obtain an overall CCR (correct classification rate) of 0.72 and a Cohen's kappa of 0.65. Comparing the predictive power of the different modeling techniques, we find that ensemble modeling (i.e. combining the projections of different individual modeling techniques) generally performs better than individual modeling techniques.

1. Introduction

Vegetation mapping is often of prime importance in monitoring, protection and restoration programs, and has therefore been the focus of many remote sensing applications (Xie et al., 2008). One particular aspect of vegetation mapping is the identification and classification of individual tree species. Being able to identify and geo-localize individual species provides useful data not only for forest management and long-term forest monitoring (e.g. Lorenz, 1995; Zweifel et al., 2010; Graf Pannatier et al., 2010; Dobbertin, 2009; Thimonier et al., 2010), but also for other application such as biodiversity assessment and conservation planning (Ferrier, 2002; Kukkala and Moilanen, 2013), or ozone risk assessment (Emberson et al., 2001; Büker et al., 2011). This latter is because sensitivity to tropospheric ozone is not only dependent on ozone exposure, microclimate and site conditions, but also on species specific, physiological and morphological, characteristics (e.g. Schaub et al., 2005). Therefore, high spatial resolution vegetation mapping is of prime importance for resource management and conservation. More generally, the mapping of individual tree species at high resolution is the first stepping-stone for making use of species-specific data such as carbon- and water-balance, sap-flow, micro-climate or transpiration in spatially explicit model parameterization, calibration, and validation (e.g. Etzold et al., 2011).

During the 1980s and 1990s, mapping and classification of individual tree species was based on the interpretation and mapping of aerial photographs. Methods have also been developed to identify individual tree crowns (Wulder, 1998). Over the last decade, high spatial resolution data (pixel size of one meter or less) became increasingly available, opening a new round of research on classifying tree species at the individual tree level (Brandtberg, 2002; Key et al., 2001; Erikson, 2004). Such information is increasingly needed to assess biodiversity effects and to map ecosystem services or landscape functions. Digital airborne data have facilitated new opportunities for tree species classification since the digital devices are considered to be spectrally and radiometrically superior to analogue cameras (Petrie and Walker, 2007). The data are recorded by frame-based sensors, e.g. Z/I DMC (Olofsson et al., 2006), Ultracam (Hirschmugl et al., 2007) or line-scanning sensors, e.g. ADS40/ADS80 (Waser et al., 2010, 2011; Waser, 2012). Complementary to spectral imagery, high-resolution airborne laser scanning (ALS) has become an

* Corresponding author. Tel.: +41 (0) 44 7392 448; fax: +41 (0) 44 7392 215.
E-mail address: Robin.Engler@gmail.com (R. Engler).

0378-1127/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.foreco.2013.07.059

operational tool in the last decade, which has rendered the classification of individual species more feasible (Brandtberg, 2007; Ørka et al., 2009; Heinzel and Koch, 2011). Airborne hyperspectral imagery has also been found to produce high accuracies for identifying individual tree species (e.g. Aspinal, 2002; Leckie et al., 2003, 2005; Boschetti et al., 2007; Jones et al., 2010). For instance, in a recent study, Dalponte et al. (2013) obtained up to 96% producer accuracies when mapping conifers in Norwegian boreal forests.

An alternative to the use of remote-sensing based variables for carrying-out projections of species distributions is the use of topo-climatic variables (e.g. temperature, rainfall, slope, solar radiation, soil water balance). Such variables have been used abundantly in the field of species distribution modeling, where spatial projections of the ecological niche of species (i.e., habitat suitability conditions) are constructed by statistically relating species observations (presence/absence or abundance) to environmental predictor variables (see Guisan and Zimmermann, 2000; Guisan and Thuiller, 2005 for a review). However, species distribution models have generally not been used to map individual trees but rather to predict habitat suitability conditions at a given site (e.g. Engler et al., 2004; Randin et al., 2006). To date, relatively few efforts have been made to combine remote-sensing data together with topo-climatic variables in order to map vegetation (e.g. Zimmermann et al., 2007), and none to our knowledge has tried to combine these data-sources for mapping the distribution of individual tree species at very high spatial resolution.

The technique of species distribution modeling has far advanced over the last two decades (Zimmermann et al., 2010), specifically with regards to the use of advanced statistical methods and the implementation of ensemble modeling techniques. Ensemble modeling (or ensemble forecasting; Araújo and New, 2007) consists in combining the projections from several different statistical modeling techniques into a single projection. The idea behind ensemble modeling is that a combined projection will have lower mean error than any of its individual constituents (Thuiller et al., 2009). Ensemble modeling has become increasingly used in bioclimatic modeling of species distribution (e.g. Engler et al., 2011), but has, to our knowledge, never been applied in a remote-sensing mapping application so far.

Here we introduce a novel approach for mapping individual tree species by using ensemble modeling and the combination of high-resolution remote-sensing data (aerial imagery) and topo-climatic variables. We map the distributions of six tree species (3 broadleaf and 3 conifers), over a 20 × 10 km study area of complex topography in North-Eastern Switzerland, using statistical modeling in conjunction with high-resolution aerial imagery (50 cm pixel size) and topo-climatic variables (5 m pixel size). More precisely, we combine six different statistical modeling techniques in an ensemble modeling approach to obtain probabilistic distribution maps for each of our target species. We additionally compare the relative predictive power of both topo-climatic and remote-sensing variables for mapping the spatial tree patterns, and we assess the importance of calibration data quality on model performance. Finally, we combine the spatial patterns obtained for each of our six individual species to obtain a map that classifies each 5 × 5 m pixel located within a forest into one of our target species, or "other" when none of our species was predicted with high enough likelihood. This allows us to map individual species for a range of purposes, or to simply re-classifying the species map e.g. into maps of plant function types (such as broadleaves and needleleaves).

2. Methods

2.1. Study area and target species

The study area is located in North-Eastern Switzerland and covers a region of 20 × 10 km centered on 9.35°E 47.36°N (Fig. 3 and Supplementary material Fig. S1). Elevation ranges from 580 to 1300 m a.s.l., and the general topography of the area is very hilly (83% of the study area has a slope > 10%). The land cover is a heterogeneous mixture of forest, grasslands, pastures, agricultural and urban areas.

The forested lands form a mosaic of fragmented surfaces that cover 27.8% of the study area (52.6 km^2; Supplementary material Fig. S1). They are mostly characterized by mixed forests with a dominance of deciduous trees along rivers and coniferous trees above 1200 m. The forests are partly managed: clearings and both deforestation and afforestation occur in several parts of the area.

For our mapping exercise we selected six tree species that are typical for the study area: *Fagus sylvatica* L., *Fraxinus excelsior* L., *Acer* sp. (mostly *Acer pseudoplatanus* L.), *Larix decidua* Mill., *Picea abies* L. and *Abies alba* Mill. The chosen species represent a balanced mix of different tree types (i.e. broadleaf vs. conifers) and occurrence frequency, with species being respectively very common (*Picea*, *Fagus*), relatively frequent (*Fraxinus*, *Acer*, *Abies*) and relatively rare (*Larix*).

2.2. Data sampling

During the summers of 2010 and 2011, 812 individual trees (130 *Fagus*, 119 *Fraxinus*, 89 *Acer*, 69 *Larix*, 271 *Picea*, 134 *Abies*) were mapped following a random-stratified sampling design. The stratification factors used in the random sampling were growing degree days above 5 °C and soil water balance (see Table 1), and ensured the collected data spanned the major environmental gradients found in our study area in a representative manner (Supplementary material Fig. S1).

Field work was carried out as follows: each visited sampling location was prospected for our six target species within a radius of about 100 m. When a species was found present at a given sampling location, one easy-to-identify individual (typically a large tree or a tree within a large patch of trees of the same species) was mapped using prints of the 50 cm resolution aerial imagery and differential GPS (±2 m accuracy). We only mapped those individuals that could be identified with very high certainty on the aerial imagery (typically large canopy trees). The field-collected data were later entered in a GIS, allowing us to manually delineate the crown of each recorded tree individual with high accuracy.

2.3. Topo-climatic variable preparation

12 different topo-climatic variables were prepared at a 5 × 5 m spatial resolution (Table 1). Slope and topographic position were computed from a 2 m resolution digital elevation model aggregated to 5 m using bilinear interpolation. Distance to the nearest water body was derived from the numeric land use model of the Swiss Federal Office of Topography (swisstopo – VECTOR25 model). All other topo-climatic variables were downscaled from existing 25 m resolution data (see Zimmermann and Kienast, 1999; Zimmermann et al., 2007 for details on computation), using bilinear interpolation. The original 25 m resolution data were computed from long-term (1961–1990) monthly means for average temperature (°C) and sum of precipitation (mm) provided for different elevations by the Swiss Federal Office of Meteorology and Climatology (MeteoSwiss), a digital elevation model and a soil suitability map (BfS, 2000).