Remote Sensing of Environment 201 (2017) 12–23

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Mapping plant species in mixed grassland communities using close range imaging spectroscopy

Javier Lopatin*, Fabian E. Fassnacht, Teja Kattenborn, Sebastian Schmidtlein

Institute of Geography and Geoecology, Karlsruhe Institute of Technology (KIT), Kaiserstraße 12, 76131 Karlsruhe, Germany

ARTICLE INFO

Keywords:
Very high resolution
Hyperspectral
AISA +
UAV
Pixel size
Diversity levels
Species cover
Functional types
Herbaceous

ABSTRACT

Grasslands are one of the ecosystems that have been strongly affected by anthropogenic impacts. The state-of-the-art in monitoring changes in grassland species composition is to conduct repeated plot-based vegetation surveys that assess the occurrence and cover of plants. These plot-based surveys are typically limited to comparably small areas and the quality of the cover estimates depends strongly on the experience and performance of the surveyors. Here, we investigate the possibility of a semi-automated, image-based method for cover estimates, by analyzing the applicability of very high spatial resolution hyperspectral data to classify grassland species at the level of individuals. This individual-oriented approach is seen as an alternative to community-oriented remote sensing depicting canopy reflectance as the total of mixed species reflectance. An AISA + imaging spectrometer mounted on a scaffold was used to scan 1 m^2 grassland plots and assess the impact of four sources of variation on the predicted species cover: (1) the spatial resolution of the scans, (2) complexity, i.e. species number and structural diversity, (3) the species cover and (4) the share of functional types (graminoids and forbs). Classifications were conducted using a support vector machine classification with a linear kernel, obtaining a median Kappa of $\sim$0.8. Species cover estimations reached median r^2 and root mean square errors (RMSE) of $\sim$0.6 and $\sim$6.2% respectively. We found that the spatial resolution and diversity level (mainly structural diversity) were the most important sources of variation affecting the performance of the proposed approach. A spatial resolution below 1 cm produced relatively good models for estimating species-specific coverages ($r^2 = \sim0.6$; RMSE = $\sim$7.5%) while predictions using pixel sizes over that threshold failed in this individual-oriented approach ($r^2 = \sim0.17$; RMSE = $\sim$20.7%). Areas with low inter-species overlap were better suited than areas with frequent inter-species overlap. We conclude that the application of very high resolution hyperspectral remote sensing in environments with low structural heterogeneity is suited for individual-oriented mapping of grassland plant species.

1. Introduction

During the last century, ecosystems have undergone an accelerated rate of environmental change due to anthropogenic impact (Smart et al., 2006). Among other impacts, these changes have affected the structure and functional composition of grasslands (Dallimer et al., 2009). Recent investigations assert that agricultural intensification – caused mainly by nutrient inputs – is one of the principal drivers of pronounced changes in grassland communities, often with associated losses in taxonomical and functional diversity (Wesche et al., 2012). Therefore, an accurate assessment of recent vegetation changes is crucial to understand current and future ecosystem dynamics. To assess these changes, two main approaches have been used in vegetation science: the establishment of permanent plots (where several measurements are repeated over time in the same plots), and the use of distribution data (where different plots are used over time, and the changes are assessed by modeling), with the first being the more reliable approach (Jandt et al., 2011).

One drawback of plot surveys is that they are expensive and time consuming when applied either to large areas or repetitively (Olsen et al., 1999). The effort and cost associated with plot surveys depend on the applied sampling approach. Most sampling approaches consider the presence and the abundance or cover of the species (Vittoz and Guisan, 2007). The cover of species can be estimated visually, by exhaustive methods or by a mixture of both. Choosing an appropriate method for cover estimation is often a trade-off between monitoring small plots with accurate methods (e.g., the pin-point method of Levy and Madden, 1933) and monitoring larger plots with lower accuracy (e.g., visual interpretation; Vittoz and Guisan, 2007). Exhaustive methods are usually very time consuming and only allow the monitoring of small

* Corresponding author.
E-mail address: javier.lopatin@kit.edu (J. Lopatin).

http://dx.doi.org/10.1016/j.rse.2017.08.031
Received 14 January 2017; Received in revised form 19 July 2017; Accepted 21 August 2017
Available online 05 September 2017
0034-4257/

vegetation plots, while visual interpretation is less time-consuming but may introduce an unknown level of observation bias into the measurements (Klimeš et al., 2001; Vittoz and Guisan, 2007).

Further problems of current vegetation survey approaches relate to the completeness of surveys and the consistency between surveyors. Vittoz and Guisan (2007) found that during vegetation surveys only about 45–63% of the species were seen by all observers (with the majority of overlooked species occurring with covers $< 0.1\%$). The study also reported that a pair of observers overlook about 10–20% less species than a single observer. The consistency between surveys also relates to the plot size as reported by Klimeš et al. (2001), who found that in larger plots the discrepancy between observers varies less (~10–20%) than in smaller plots (~33%). As species cover is more similarly estimated in small plots than in larger ones (Sykes et al., 1983), the use of frames that include a known number of grid squares can also increase the similarity of estimates among observers. This may even result in accuracies more similar than the pin-point method with trained observers, as well as requiring only half the field time (Sykes et al., 1983). In summary, the results of current vegetation surveys are likely to vary with the experience of the surveyor and contain a notable degree of uncertainty due to potentially missed species. The development of new, user-independent and objective methods that combine the high level of detail of exhaustive methods with high accuracies and time-efficiency (and eventually also with an option to cover larger areas), would be a great advance for vegetation monitoring.

Automated remote sensing (RS) techniques have been applied to classify individual trees since the 1980s (see Fassnacht et al., 2016 for a comprehensive review), while for grassland species few efforts can be found in the literature. In grasslands, we differentiate between community-oriented approaches and individual-oriented remote sensing approaches. Community-oriented approaches treat canopy reflectance as an expression of mixed species reflectance. In gradient-based analyses, individual grassland species occurrences can be retrieved in a top-down approach relying on the occurrence of species along mapped gradients (Neumann et al., 2016, 2015). While this is suited for situations where species are hiding in sub-pixel information, individual-oriented methods are addressing species separately. Following the latter strategy, Gebhardt et al. (2006) used very high spatial resolution data (~0.6 mm) from an RGB camera in an experimental setup to classify *Rumex obtusifolius* with an object-oriented approach. Similarly, Silva et al. (2014) used RGB imagery (~1 cm pixel size) collected from a balloon-mounted camera to classify two species (*Setaria sphacelata* and *Pteridium arachnoideum*) using texture and object-based information. Booth and Cox (2008) used RGB imagery to estimate cow manure, green-grass and shrub covers under different grazing intensities in Colorado, USA. Further examples include the studies of Kumar and Sinha (2014), who successfully classified four species of salt-marsh vegetation in Australia using Quickbird data, and the studies of Andrew and Ustin (2008) and Lu et al. (2009), who used aerial hyperspectral data to identify *Lepidium latifolium* in riversides in California.

Nevertheless, no attempt has been made so far to classify all grassland species in a given area using hyperspectral imagery with very high spatial resolution (< 1 cm). In case of success, such an approach could be a turning point towards RS-based semi-automatic grassland surveys, covering reasonably large areas while limiting field-work to few calibration and validation plots. Unmanned aerial vehicles (UAV) arise as a suitable option for such a monitoring approach, due to e.g. their high spatial resolution and relatively low acquisition costs. Unfortunately, UAV-based sensors have not yet reached the ideal geometric or radiometric quality (e.g., close to the radiometric resolution of airborne spectrometers) for this task. Nevertheless, we believe that in the future small sensors with appropriate spatial and radiometric resolution will become available for UAV RS.

In this study, we classified grassland species using an AISA + Eagle imaging spectrometer mounted on a scaffold at a height of 2.5 m above ground, in order to simulate future UAV-based image qualities. We collected hyperspectral images of one square meter field plots and subsequently classified all plants to obtain cover estimates of each species.

The aims of the study were to assess the feasibility of this approach for classifying grassland species, and to determine under which conditions this method could be useful in practice. To address these questions, we analyzed the influence of four sources of variation on the obtained results: the spatial resolution of the images (leading to an increased proportion of mixed pixels), the complexity of the grassland (along a gradient in species numbers and structural diversity), the species-specific cover values and the functional type of the species. We hypothesized that:

(1) As the spatial resolution decreases, the agreement between field-estimated and remotely-sensed covers will decrease due to an increased proportion of mixed pixels;
(2) While the complexity gradient increases, the correlation between field-estimated and remotely-sensed cover values will decrease due to an increased number of species occurring in the understorey;
(3) Species with higher cover will have a higher chance to be classified accurately; and
(4) Forbs will be classified with greater accuracy than grasses due to their broader leaves (higher chances of clearly assignable leaf spectra).

In addition, we examined the role of the classifier and the method for defining training areas in the hyperspectral scans as potentially influential factors.

2. Material and methods

2.1. Study area and field data

The study areas were grassland patches located in the Botanical Garden and other sites belonging to the Karlsruhe Institute of Technology, Karlsruhe, Germany. We selected four areas with different management treatments, including parks and an abandoned construction site. The treatments, including annual mowing and water irrigation, led to a species and structural diversity gradient which was suitable to test our hypotheses. Field plots were randomly placed inside those areas, keeping a minimum distance of 10 m between plots (Fig. 1E).

In August 2016 a one square meter sampling frame was used to sample the grassland species (Fig. 1B i) within 11 plots. The frame contained 16 sub-plots of 0.25 m × 0.25 m, where the species' positions (Fig. 1B ii) and covers were recorded using the exhaustive pin-point (or point intercept) method with a systematic grid of 5 cm, and a total of 25 points per sub-plot (Fig. 1B iii). The survey resulted in a total of 176 sub-plots.

We sampled a total of 42 species (forbs = 36, graminoids = 5 and bryophyte = 1; see Appendix A), with a maximum of 15 species per sub-plot (Table 1).

2.2. Complexity gradient

Diversity levels were defined for each selected study areas according to their species and structural diversity. Species diversity was determined by species richness and evenness – defined by Camargo (1992): $evenness = richness - \Sigma_{i=1}^{n} |P_i - P_j|$, where P_{ij} is the relative abundance of a species i in a biological community h – while structural diversity was defined using the inter-species canopy mixture level (see Table 1, Fig. 1D). We will refer to this interaction between species and structural diversity as the complexity gradient. The four defined complexity levels are:

- *Complexity 1*: This category is characterized by high species richness