Application of hyperspectral remote sensing for flower mapping in African savannas

Tobias Landmann a,⁎, Rami Piiroinen b, David M. Makori a,d, Elfati M. Abdel-Rahman a,c, Sospeter Makau a, Petri Pellikka b, Suresh K. Raina a

a International Center for Insect Physiology and Ecology (ICIPE), P.O. Box 30772, 00100 Nairobi, Kenya
b Department of Geosciences and Geography, University of Helsinki, Gustaf Hallströminkatu 2, Helsinki, Finland
Department of Agronomy, Faculty of Agriculture, University of Khartoum, Khartoum North 13314, Sudan
d Department of Geography, School of Agriculture, Environment and Earth Sciences, University of Kwazulu Natal (UKZN), 1 Curbi Road, Pietermaritzburg 3209, South Africa

A R T I C L E   I N F O

Article history:
Received 12 February 2015
Received in revised form 22 May 2015
Accepted 4 June 2015
Available online xxxx

Keywords:
Airborne hyperspectral data
Linear unmixing
Change Vector Analysis
Phenology
Flower mapping

A B S T R A C T

We tested the suitability and accuracy of hyperspectral data to produce the first African flowering and short-term floral cycle map. The spatial distribution and abundance, as well as the floral cycle, of melliferous plants are of utmost importance for evaluating pollination effects and to understand the relationship between melliferous plants in the landscape and the quantity and quality of bee keeping products. For a study site in Kenya, airborne AISA/Eagle hyperspectral data with 60 cm pixel resolution (400 to 990 nm spectral ranges) was captured in January 2014, at the beginning of the prime flowering period, and during the prime flowering period in February 2013. Aerial digital imagery with 10 cm pixel size and Smartphone captures in the field were used for reference data collection. The flowering species were grouped into functional flowering plant groups. Linear spectral unmixing and Change Vector Analysis (CVA) were used on the bi-temporal AISA/Eagle data to produce a hard cover map showing the spatial distribution, abundance and short-term flowering cycle of melliferous plants. Overall accuracies were slightly higher in the February 2013 imagery at the prime flowering period; all flowering plant groups together ("All") could be mapped with an overall accuracy of 83% (n = 512). The "White forbs" flowering plant group was most accurately mapped in both AISA/Eagle acquisition dates. Based on Duncan’s inter-class similarity test, the "White forbs" group was also most distinct from other flowering plant groups. There is a need to investigate the effect of spectral endmember variability and upscaling options for space-borne monitoring of the floral cycle at key sites in Africa. Floral cycle maps can help decision makers and bee keepers to understand how bee colonies interact with the floral environment and what to expect from an apiary in terms of honey flow.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

In Africa a significant part of the rural population draws their income from beekeeping (Qaiser, Tahir, Taj, & Ali, 2013), however information about the relationship between the abundance and distribution of flowering plants in the landscape and hive productivity is largely unavailable (Evans & Schwarz, 2011). The floral cycle and flowering intensity in semi-arid Africa is spatially and temporally highly variable depending on the plant species as well as land form, and local climate and edaphic conditions (Raina & Kimbu, 2005). The floral cycle refers to the duration of the blossoming period and flowering intensity which includes the fractional coverage of flowering buds within a single tree, plant or vegetation community (Mcintosh, 2002). In situ observations of the floral cycle of key melliferous plants are often used to produce floral calendars for a specific site (Raina, Kioko, Zethner, & Wren, 2011). Floral calendars categorize various flowers, their value to bees, abundance, season and duration of bloom for a given area. Melliferous plants produce nectar and pollen which is collected by honey bees and converted into honey (Decourtye, Mader, & Desneux, 2010). Spatio-temporal information about the distribution and abundance of flowering plants can help to understand bee foraging behavior (Smith, Lopez Quintero, Moreno Patino, Roubik, & Wcislo, 2012) and ultimately, if a link can be established between ecosystem integrity and hive productivity, conservation incentives can be formulated (Abou-Shaara, 2014). Ecosystem integrity is a term used to describe the degree of self-maintenance of ecosystem functions within a particular ecosystem (Haines-Young & Potschin, 2010). A decrease in the fractional cover of melliferous plant communities within agro-ecological landscapes, due to for instance deforestation, would mean that ecosystem integrity in view of pollination and beekeeping is compromised. Lateral to that, long-term spatial data on landscape flowering patterns would also provide a critical understanding about nutritional, climate and
ecological stresses which trigger pest and diseases in bee colonies (Aizen & Harder, 2009).

Earth observation (EO) is an effective tool to spatially assess plant traits, vegetation condition and ecological processes (Chu & Skidmore, 2006; Rocchini, 2010). Well validated remotely sensed data sets are more cost and time effective compared to sampling the same area using field based approaches (Weng, 2002). Moreover, when several well calibrated satellite images are used in time-series, phenological stages including the floral cycle can be effectively mapped (Ge, Everitt, Carruthers, Gong, & Anderson, 2006). EO mapping routines still have a vast unexplored potential to synoptically capture landscape ecological parameters in relation to pollinators and the distribution, temporal development and status of melliferous plants (Chen, Shen, Zhu, & Tang, 2009).

In airborne hyperspectral remote sensing dozens to hundreds of spectral bands are commonly used for the qualitative and quantitative analyses of various materials and targets across the visible (VIS), near-infrared (NIR) and short-wave infrared (SWIR) spectral ranges (Ustin, Roberts, Gamon, Asner, & Green, 2004; van der Meer et al., 2012). With its high spatial and spectral resolution, hyperspectral remote sensing can help map vegetation communities down to the species level, and detect elementary, biogeochemical and physiological vegetation properties such as cellulose and lignin content (Ustin et al., 2004). However, the large data volumes and high acquisition costs of specifically airborne hyperspectral data can be challenging to institutions and researchers alike (Higgins, Asner, Perez, Elespuru, & Alonso, 2014). Hyperspectral remote sensing data interpretation is often hampered by large data dimensionality, unavailability of multi-temporal and time-synchronous and, if spectral endmembers (EM) are used within a linear mixture modeling approach, high variability of spectral EM in time and space (Somers, Asner, Tits, & Coppin, 2011). To the best of our knowledge there are currently no remote sensing studies in Africa that specifically aim at mapping the abundance, distribution and the short-term flowering cycle of melliferous plants on a landscape scale. Some earlier studies in North America mapped the floral cycle and phenological characteristics of specific invasive weeds within the landscape mosaic using airborne and ground based hyperspectral data (Ge et al., 2006). Invasive species often exhibit a distinct floral cycle, which enhances their spectral differentiation at certain periods within the phenological cycle (Lawrence, Wood, & Sheley, 2006; Parker Williams & Hunt, 2002). Carvalho, Schler, van der Putten, and Skidmore (2013) used hyperspectral-based spectroscopy to assess the flowering and leaf phenology of Jacobaea vulgaris in northern Europe as a means of obtaining an insight into the chemical characteristics of the plant. Chen et al. (2009) used hyperspectral data and spectral mixture modeling to develop an accurate flowering coverage for Halterpestes tricuspis in Tibetan plateau grasslands. The authors found that yellow flowering H. tricuspis were best mapped using the 500–670 nm spectral region (visible to near-infrared wavebands) while no apparent flowering effect was found in the NIR spectral region.

Shen, Chen, Zhu, Tang, and Chen (2010) found that flowering in the same species (H. tricuspis) significantly affects the robustness of the Enhanced Vegetation Index (EVI) which is often used in wide-area vegetation monitoring studies. In a study in European orchards, hyperspectral data and waveband selection was employed to detect the spatial abundance and distribution of floral pear buds (Wouters, De Ketelaere, De Baerdemaeker, & Saëys, 2013). In nearly all studies reviewed, in situ hyperspectral data sets at plot level, from hand held devices, were used for flower or floral bud mapping and in most cases only a single observation period was considered. In this study multi-temporal (bi-temporal) AISA/Eagle hyperspectral aircraft data was used as input. Two commonly used spectral mapping algorithms were used in sequence, namely linear spectral unmixing and Change Vector Analysis (CVA) (Dubovyk, Menz, Conrad, Thonfeld, & Khazmiza, 2013). Linear spectral unmixing is widely used on hyperspectral data for landscape feature mapping (Clark & Roush, 1984) mainly due to the relative simplicity of the spectral mixture model used and easy interpretation of results (Dobigeon et al., 2014). Linear spectral unmixing results are considered to be robust if well defined and stable reference spectra (or spectral endmembers – EM) are utilized in the model approximation. Linear spectral unmixing has some uncertainties mainly due to its inability to accurately consider non-linear spectral mixing effects (Altmann, Dobigeon, McLaughlin, & Tourneret, 2013) within highly heterogeneous landscapes (Higgins et al., 2014; Somers et al., 2011). CVA was, thus, utilized to fine-tune the attained linear unmixing results. In CVA the two-dimensional distance (magnitude) and direction between spectra or spectral derivatives is computed (Chen, Gong, He, Pu, & Shi, 2003). Albeit the difficulties in interpreting CVA results, CVA is considered a state-of-the-art change detection algorithm that is able to accurately map fine scale spectral changes in multi-temporal imagery (Landmann, Schramm, Huettich, & Dech, 2012).

2. Methodology

2.1. Study site

The study site is located in the Mwingi Central Sub County, Kitui County in Kenya (Fig. 1). It covers an area of 100 km². Mwingi is a semi-arid area with two rainfall peaks in April (147 Mean Annual Precipitation) and November (270 Mean Annual Precipitation). The average maximum temperature is 31 °C while the minimum is 15 °C. The hottest months are February–March and September–October, with the coldest months being July to August. The main woody species in Mwingi include Acacia spp., Melia volkensii, Azadirachta indica, Zizyphus abyssinica, Albizia gummifera and Markhamia lutea. The main flowering plants include Acacia spp., Terminalia brownie, Aspilia mozambicens, Cassia diabotica, Cassia sennea, Solanum incumum, and Boscia and Creova spp. The main flowering period for most of these plants is from January to May (Raina & Kimbu, 2005) with a few plants flowering in December. Relatively high rainfall amounts in November to December trigger the flowering of most plants from January onwards. This flowering trend is sustained by further rainfall during March and April.

The study area is mostly an agro-ecological mosaic, whereas the main crops are maize (Zea mays) and sorghum (Sorghum bicolor), and additional income is generated from beekeeping. Deforestation of the few patches of near-natural vegetation leads to the reduction of the diversity of melliferous plants and consequently reduction in honey production (Delaplane, 2010). The main flowering species are Acacia spp. which bloom from February to April with a pronounced flowering peak in February, while the two main crops (maize and sorghum) have their main flowering period in January (Nagarajan, Audi, Jones, & Smale, 2007).

2.2. Hyperspectral data acquisition and pre-processing

AISA/Eagle imaging spectrometer (Specim Limited., Finland) was used to cover the study area with hyperspectral data from an airborne platform using a maximum flight altitude of 860 m above ground level. AISA/Eagle is a pushbroom scanner with instantaneous field of view of 0.037°, field of view of 36.04° and 969 pixels across the spatial axis. The sensor was used in 8 times spectral binning mode, which produces output images in 64 bands with a full width at half maximum (FWHM) of 8–10.5 nm in the spectral range of 400–1000 nm. Spatial resolution was 0.6 m after geo-referencing. The flight campaign in year 2013 carried out on the 14th of February during the maximum flowering period and in year 2014 on the 11th of January at the beginning of the main floral period. Fig. 2 shows the various input data sets used in this study as dashed rectangles, processing routines as rounded rectangles and derived data sets as solid line rectangles. Radiometric calibration and geolocating was performed using the CaliGeoPro tool (Pre-processing in Fig. 2) (Specim Limited., Finland).
Fig. 1. The red rectangle shows the Mwingi study area (100 km²) which lies north-east of the Kenyan capital Nairobi. The study area lies within the Mwingi Central Sub County, within the Kitui County. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 2. Data flow diagram showing input data as dashed rectangles, main data processing steps are rounded rectangles and square rectangles illustrate preliminary (derived) data sets. Arrows show data flow directions and tie-up points (PCA — Principle Component Analysis; CVA — Change Vector Analysis; fl. — flowering).
Direct georeferencing was used. Herein pixel positions are calculated from the position and orientation of the sensor and georeferenced on a coordinate plane with elevation values derived from digital elevation model (DEM). The used DEM was derived from global C-band shuttle radar topographic mission (SRTM) DEM with 90 m pixel resolution and interpolated to match the target resolution of 0.6 m (Rabus, Eineder, Roth, & Bämler, 2003). The root mean square error (RMSE) for the SRTM data was estimated to be 8.6 m although higher errors have been observed especially in areas with steep slopes and forest cover (Shortridge & Messina, 2011). ATCOR-4 was used for atmospheric water vapor and haze reduction (Richter & Schläpfer, 2002). As ground reference spectra or field atmospheric data were not available, we relied on automatic JUT retrieval and modeling of atmospheric parameters, which are based on the MODTRAN (MODerate resolution atmospheric TRANsmission) code. We mosaicked the georeferenced flight lines together using nearest neighbor resampling. The ‘Georeferenced Mosaicking’ method used did not change the pixel spatial properties.

It was established from collected field data that there was a need to further geolocate the AISA/Eagle imagery (after mosaicking the data strips) since the established RMSE error of 8.6 m for the SRTM orthorectification was too high. A WorldView-2 image with a pixel size of 2 m captured in April 2014 over the Mwingi study site was used for final co-registration of the AISA/Eagle data mosaics using nearest neighbor interpolation. This working step is part of ‘pre-processing’ in Fig. 2.

2.3. Reference data collection

Field sampling points were collected during the flight campaigns in February 2013 and in January 2014 using a straightforward stratified random sampling method. In total over 100 random sample points were initially generated for both dates (2013 and 2014 sampling) using the Geospatial Modeling Environment software. A Global Positioning System (GPS) device with an accuracy of 3 m was used to locate the generated points in the field. Flowering plants closest to these points were geotagged using Smartphones. Flowering crops (maize and sorghum) were only sampled in January, since this is the main crop flowering period in which both crops generally have a high flowering compaction and also a high background ‘green’ leaf coverage. Name of the flowering species as well as the canopy size (for trees and shrubs), flowering color and percentage (visually estimated) were recorded for each plant. Geotagged photographs of each plant were taken from the main four cardinal directions. The Smartphone-based geotagged photos were later combined with the field records to create a digital and comprehensive reference data set on flowering plants (referred to as ‘reference data set’ in Fig. 2). For the 2013 data, additional flowering plant species were added from 10 cm visual imagery captured from a NIKON D3X digital camera which was simultaneously flown during the 2013 airborne campaign (Fig. 2). The digital camera images were captured in a 5:4 image format to reduce unnecessarily large side overlaps since the flight lines were planned according to the AISA/Eagle field of view parameters. This resulted in a side overlap of 67% which was adequate for mosaicking (and geo-rectification) using the EnsoMosaic tool from MosaicMill.

In January 2014 (beginning of the dry season), ‘brown’ (non-chlorophyll active) trees were found scattered throughout the landscape, so over 25 reference plants were sampled for “Brown” in January 2014. In total there were 146 reference data points available for 2013 and 2014, respectively, and in total 19 different flowering plant species could be determined for both dates. Since canopy level flowering reflectance is largely affected by background plant characteristics (Gai, Fan, Xu, & Zhang, 2011; Shen et al., 2010) and mapping flowering in individual species was not the aim, the flowering species were grouped into functional flowering plant groups (Table 1) (Fig. 2). The groupings (classes) were formed according to whether flowering occurs within a ‘green’ canopy or not; color of the flowers (white or yellow) and plant morphology. Table 1 shows the functional flowering plant groups; the dominant species within that group; the plant morphology; background leaf and flowering color and number of reference pixels.

After overlaying the number of samples on the AISA/Eagle imagery for 2013 and 2014, respectively, more pixels, identified as belonging to the same tree or shrub, were sampled on the image. The final collect (last column in Table 1) was then used for the accuracy assessment described in Section 2.5.

2.4. Development of the flowering mapping methodology

The flowering mapping approach used herein was based on the spectral properties of flowering plants, separability of flowering plants from other spectral features, intrinsic data dimensionality and intra-annual phenological behavior of melliferous plants at the Mwingi site (Raina et al., 2011). Fig. 3 shows a good spectral separability between the EM (that is spectral signatures) representing ‘green’ that is chlorophyll active vegetation and flowering plants in the 550 nm to 680 nm spectral region (VIS wavebands). A subset of the field data was used to extract the representative spectral EM in Fig. 3. The NIR region exhibited a large EM variability for ‘flowering’ and thus, if the NIR region was to be used in isolation for floral mapping, high spectral confusion would occur. Increased spectral variability causes overlapping spectral responses that would create confusion between classes to be mapped.

2.4.1. Spectral linear unmixing and endmember selection

An unconstraint linear spectral unmixing mapping approach was used. In the unconstrained method, abundances may assume negative values and are not constrained to sum-to-unity (one) (Yang, Everitt, & Du, 2010). We opted for linear unmixing since (i) linear unmixing results are widely considered to yield robust results if EM are known from reference data (Schramm, Landmann, Lohmann, & Heipke, 2008), and (ii) a non-linear unmixing model would assume non-linear and complex relationships between pixels making results difficult to interpret (Borel & Gerstl, 1994). Moreover, it was aimed to assimilate the EM abundances for flowering and ‘green’ into the CVA computation to reduce possible over-exaggeration of the EM abundances in the linear unmixing model. Fig. 2 shows these sequential working steps. The

Table 1
Overview of the reference functional groups derived from the reference data collected for February 2013 and January 2014. The functional group allocations are shown in the first column, the second column shows the dominant species in that group, and the other columns show plant group characteristics. Number of samples refers to the number of reference plant species sampled for each of the respective observation periods. Values in parenthesis represent the total number of reference pixels used (final collect) after additional pixels were collected on the imagery (n/a — not applicable).

<table>
<thead>
<tr>
<th>Functional group</th>
<th>Dominant species</th>
<th>Color of flower/background leaf</th>
<th>Number of samples (final collect)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>January 2014</td>
</tr>
<tr>
<td>Brown leaves</td>
<td>n/a</td>
<td>n/a</td>
<td>27 (0)</td>
</tr>
<tr>
<td>Crops</td>
<td>Zea mays/Sorghum bicolor</td>
<td>White/green</td>
<td>14 (84)</td>
</tr>
<tr>
<td>White acacia</td>
<td>Acacia tortilis</td>
<td>White/n/a</td>
<td>5 (151)</td>
</tr>
<tr>
<td>White forbs</td>
<td>Ipomoea kitaenisis vatke</td>
<td>White/green</td>
<td>14 (73)</td>
</tr>
<tr>
<td>White green</td>
<td>Terminalia brownee</td>
<td>White/green</td>
<td>57 (315)</td>
</tr>
<tr>
<td>Yellow green</td>
<td>Acacia nilotica</td>
<td>Yellow/green</td>
<td>43 (129)</td>
</tr>
</tbody>
</table>
multi-temporal CVA approach was also employed to reduce spectral confusion between 'brown' and flowering plants (Fig. 3).

The number and representativeness of EM that define the spectral sub-space is of utmost importance in linear unmixing modeling (Adams et al., 1995). As the number of EM and their representativeness is often related to the data dimensionality, a Principle Component Analysis (PCA) on the 64 band AISA/Eagle imagery was carried out (denoted 'PCA' in Fig. 2). The results showed that more than 89% of the variance was due to soil reflectance, 'green' or chlorophyll active vegetation and 'other' vegetation. Flowering plants were found to make up most of the 'other' vegetation EM. From this finding and the fact that most flowering plants at the Mwingi site are 'green' in January and flowering in February (Raina et al., 2011), it was assumed that most pixels, in both AISA/Eagle data sets, contain EM fractions of soil reflectance, 'green' and 'other' vegetation that is flowering plants. The mean spectral EM for flowering was derived from a subset of the reference pixels (Table 1), while the EM means for 'bare soil' and 'green' were collected from the bi-temporal imagery itself (Dennison & Roberts, 2003).

The linear unmixing function of the EM fractions for flowering (or other vegetation), 'green' and bare soil were summed and defined for each pixel as (Asner & Heidebrecht, 2002);

\[
\rho(\lambda)_{\text{pixel}} = \sum \{Ff\rho(\lambda)e + Fg\rho(\lambda)f + Fb\rho(\lambda)b\} \tag{1}
\]

where \(\rho(\lambda)e\) is the reflectance of each EM (e) at waveband \(\lambda\), and \(F\) is the sub-pixel abundance of each feature solved for, in this case flowering (\(f\)), 'green' vegetation (\(g\)) and bare soil (\(b\)).

Pixels where the sub-pixel abundance of the flowering EM (\(Ff\)) were above 0.5 were tagged in both AISA/Eagle data sets to produce a 'preliminary' flowering abundance map. Using visual inspection, it was observed that most flowering reference pixels were mapped after the 0.5 threshold for (\(Ff\)) was applied.

2.4.2. Change Vector Analysis (CVA) for enhanced mapping

In the second processing step Change Vector Analysis (Chen et al., 2003) was employed to further reduce mapping errors due to spectral confusion of flowering plants with senescent ('brown') vegetation and areas where background soil reflectance would perturb the flowering signal. In CVA the magnitude and direction of n-dimensional change is computed using selected variables (Landmann et al., 2012). In this study the per pixel abundances of the 'green' EM (\(Fg\)) and the abundances of the flowering EM (\(Ff\)) from the unmixing result are used as input variables (Dubovyk et al., 2013). These two variables are suited to characterize the phenological cycle of most melliferous plants found within the study site (Raina et al., 2011). The assumption is that dominant vegetation communities such as Acacia tortilis and Acacia nilotica are 'green' (non-flowering) in the pre-flowering period (January data) and flower during the peak flowering period (February data). We use CVA as a change detection tool to mimic this phenological behavior of flowering plants at the study site using bi-temporal hyperspectral data.

The CVA magnitude function can be estimated as;

\[
|c| = \sqrt{(\rho(Ff, z) - \rho(Ff, y))^2 + (\rho(Fg, z) - \rho(Fg, y))^2} \tag{2}
\]

where \(\rho(Ff, y)\) and \(\rho(Ff, z)\) are the abundances of the flowering EM of pixel \(i\) at time \(y\) and \(z\), respectively, and \(\rho(Fg, y)\) and \(\rho(Fg, z)\) are the corresponding abundances of the 'green' EM of pixel \(i\) at time \(y\) and \(z\), respectively.

The CVA direction \(d\) can be estimated as an angle expressed in degrees of a full circle;

\[
d = \arctan \left( \frac{(Ff, z) - (Ff, y)}{(Fg, z) - (Fg, y)} \right) \tag{3}
\]

Here \(d\) changes the increase in per pixel EM abundances between the two dates as an angle that is full circle from 0 to 360°.

Both the magnitude and the direction were subsequently used to refine the linear unmixing result derived in formulae (1). All per pixel CVA directions (angles) between 90° and 180° derived in formulae (2) were tagged as flowering plants in the February 2013 image. This range showed a decrease in 'green' EM abundance and an increase in the 'flowering' EM between January and February, indicating that in January the plant was non-flowering ('green') and in February the same plant is flowering. Similarly, all directions between 270° and 360° were tagged as flowering plants in the January 2014 image. This range showed an increase in the flowering EM abundance between February and January, while the 'green' EM showed no change or decreased. Directions between 0° and 90° showed 'near to no change' in EM abundances of 'green' and flowering between the two observation periods. These pixels were tagged as plants that flower in both acquisition dates such as Aspila mozambiensis or Casia spp.

Regardless of the above defined CVA direction ranges, all pixels exhibiting CVA magnitudes smaller than 1.5 were excluded, i.e. mapped as non-flowering. Pixels below this threshold were found, from visual inspection of the AISA/Eagle reference data, to be largely related to 'other' vegetation phenology effects such as 'brown' leaf shedding between January and February. Some of these pixels also had high soil EM abundance at either one or both acquisition dates.
2.5. Accuracy assessment

The overall accuracy assessment was performed on all flowering groupings together for 2013 and 2014, respectively (“All”), and also for each functional flowering group and year separately (denoted ‘Accuracy assessment’ in Fig. 2).

Intra class and inter class variability measurements were, furthermore, computed in order to determine the validity, accuracy and essentially degree of separability of the flowering groupings (classes) (Tits et al., 2012). Box plot variability and the standard variations (SD) for each flowering group were used to illustrate how the intra class variability differs between the functional flowering groups. The flowering EM abundances, from the spectral unmixing result, were used as data elements for the flowering groups (Table 1). The box plots and the SD are, thus, based on the distribution of the flowering EM within a given flowering group. Per class SD is one method to determine intra class variability indicating class mapping accuracies (Foody & Doan, 2007).

Inter class variability was determined for the flowering groups using the Analysis of Variance (ANOVA) Duncan’s test for class similarity. The same flowering EM abundance distributions within one flowering group were used. Significant differences between the flowering groupings at 95% confidence levels were determined; groups with significant differences were assigned letters (a, b, c, d). Those groups that contained the same letters did not have significant differences between them.

3. Results and discussion

3.1. Spectral unmixing and CVA mapping results

The EM abundance mapping results computed in formulae (1) are shown in Fig. 4. The unmixing results for January 2014 (pre-flowering season) are shown in Fig. 4a and b showing the unmixing result for February 2013 (peak flowering season), Fig. 5 shows the final flower mapping result computed using formulae (1), (2) and (3), for the same subset, after thresholding the CVA directions and magnitudes. Both AISA/Eagle acquisitions dates (January 2014 as well as February 2013) were considered in the hard classification results illustrated in Fig. 5a and b (zoom image). The hard classification result for January 2014 showed that flowering plants made up 3.9% of the total 100 km² study area and 7.8% percent of all vegetation within the study area. In February 2013, the aerial coverage for ‘flowering’ as a function of the total area of the study site was found to be 3%, while 6.2% of all vegetation was mapped as ‘flowering’. The relatively high flowering percentages for January (pre-flowering season) can be explained by the presence of flowering maize and sorghum fields that can be seen as yellowish patches in Fig. 5a and b. The zoom image (Fig. 5b) taken from another area in the overall study site, shows flowering maize from the January data alongside flowering acacia trees from the February data relatively clearly.

The field derived flowering data sheet (Table 1) confirmed this finding, i.e. up to 35% of all flowering plants tagged in the field in January 2014 were flowering maize or sorghum. In February 2013 (peak flowering period), most flowering species using the field reference data were found to be A. tortilis and yellow flowering A. nilotica. These two Acacia species were also found predominantly scattered on lower altitude planes (below 1000 m. altitude), or clustered along seepage lines or within riverside vegetation communities.

These spatial flowering patterns and topography effects are visible in the February AISA/Eagle results (Fig. 5a and b). At the Mwingi study site, areas above 1200 m in elevation as well as mid slopes, scarps or crests are generally dominated by Combretum tree species while the lower altitude sites are mostly dominated by Acacia species (Muok, Owuor, Dawson, & Were, 2000). Combretum species flower in April and October, within the two rainy seasons, and they are less important for bee biodiversity and honey production (Raina & Kimbu, 2005).

Due to the non-availability of comparable flower mapping studies and the unique combination of spectral unmixing with CVA for feature mapping in African savannas, a comparison with other flower mapping study results in Asia or North America would not be feasible. It is preferable to aim at ascertaining site specific accuracies of the flower mapping result. The next section deals with the accuracies for each functional flowering group and uncertainties of the mapping approach in terms of EM variability and flowering characteristics.

3.2. Accuracy assessment

Table 2 shows the overall accuracy (producer’s accuracy) results for each of the AISA/Eagle mapped flowering plant groups and for all the mapped flowering groups together (‘All’). The accuracy results are shown separately for January 2014 and February 2013. Table 3 shows the confusion matrix for all the mapped flowering groups (‘All’) for the two AISA/Eagle observations, respectively. The overall accuracy for all flowering groups together (‘All’) was 80.7% (n = 752) for the January 2014 data and 83% (n = 512) for the February 2013 data. The flowering group “White forbs”, which is dominated by Ipomoea kituensis varite, showed the largest overall accuracy score in both AISA/Eagle image acquisitions. The accuracies were 98% for the January data (n = 73) and 97% for the February data (n = 77). The flowering group “White acacia”, which is dominated by A. tortilis and Acacia brevispica, had the second highest accuracy score in the February 2013 data (86%, n = 211), while “Crops” had the second highest accuracy score in the January 2014 data (96.4%, n = 84).

The confusion matrix in Table 3 showed a slightly higher user accuracy of 87.4% (n = 752) for “All” functional flowering groups (classes) mapped using the January 2014 data, while the February 2013 mapping result had a lower user accuracy score at 74.8% (n = 512). The lower user accuracy but high producer accuracy for February 2013 means that a significant number of pixels from the field data were correctly classified as ‘flowering’ (high producer accuracy), while there is a probability (proportionate to the error) that pixels classified as ‘flowering’ were actually other features (i.e. bare surfaces, scanty vegetation or green canopies) or did not even exist on the ground (low user’s accuracy). Using visual inspection it was found that some white acacias were mapped as ‘bare soil’ in the February 2013 imagery.

Intra as well as inter class (group) variabilities, as a function of the per pixel flowering spectral EM abundances from formulae (1), provided a further insight into the mapping accuracy of the flowering groups (Somers et al., 2011; Tits et al., 2012). The boxes plots in Fig. 6 illustrate the intra class variabilities for all the mapped flowering groups together in one group (‘All’) for January 2014 and February 2013, respectively. The “All” group intra variability, using the EM abundances, was generally lower in the February 2013 data (Fig. 6). This lower intra group variability (Fig. 6), paired with a slightly higher overall accuracy for the “All” group in February 2013 (Table 2), can be explained by the generally more compact flowering patterns we observed in the field in February 2013 (maximum flowering period). Shen et al. (2010) confirmed that background heterogeneity, i.e. when for instance ‘green’ leaf reflectance is mixed with the flowering signal, is a key determinant in accurate vegetation and essentially flower mapping.

Fig. 7 shows the intra and inter class EM based variability for each of the flowering groups for the two observation periods. The intra class variability of the groups is herein determined by the standard deviation (SD) for each group (Fig. 7). The inter class variation is considered by applying the Duncan’s test for class similarity to the flowering EM abundance based flowering groups (Foody & Doan, 2007).

The inter class variabilities for flowering groups in February 2013 were generally higher than in the January 2014 data. Only “White green” and “Yellow green” did not have a significant difference in their per pixel abundances according to the Duncan’s group variance test (p ≤ 0.05) (Fig. 7). The mean SD of the January 2014 flowering groups was also higher which confirms that the February 2013 flowering groups were generally more explicitly mapped.
The Duncan’s test for the January 2014 data showed that only the flowering group “White forbs” is dissimilar to the other flowering groups. “Yellow green” are similar to “White acacia” and “Crops” are similar to “White green”. The group pairings mentioned have the same Duncan’s scores (b and c, respectively) (Fig. 7). However, since the group “Crops” has a relative high overall accuracy (Table 2), this group can be considered to be mapped with an acceptable accuracy albeit high confusion rates with “White green”.

In both AISA/Eagle observations the group “White forbs” proved to be most separate from other flowering plant groups, using the Duncan’s test. This group also exhibited the highest overall accuracy scores in both AISA/Eagle data sets (Table 2). This group can therefore be

Fig. 4. AISA/Eagle based EM abundance map for January 2014 in (a) and EM abundance map for February 2013 in (b) for a 2.5 km² subset of the study area.
considered to be most accurately mappable using 60 cm AISA/Eagle hyperspectral imagery.

Flowering groups that exhibited a large fractional coverage of ‘green’ background foliage and/or are sparsely flowering such as “Yellow green” (*A. nilotica* and *Aspilia mozambensis*) or “white green” (*Terminalia brownee*) had lower overall accuracy scores (Table 2) and also high inter class variabilities (Fig. 7). In general, mapping inaccuracies in hyperspectral-based modeling and mapping can be explained by high spectral EM variability in time and space (Somers et al., 2011), spectral mixing effects and intra-species as well as inter plant physiology and

![Fig. 5. Final flowering and short-term flowering cycle map combined for both AISA/Eagle acquisition dates is shown in (a). Map (b) is a zoom in image of the final flowering map. The zoom image is taken from another part of the study area.](image-url)
phenological variations (He, Rochini, Neteler, & Nagendra, 2011). This specifically means that the canopy-level flowering spectral response, as a phenological effect in time and space, has to be larger than the leaf level spectral response. The intra versus the inter class accuracies indicated that color of flowering and general compaction of functional flowering classes plays a role in detection accuracies. Flowering compaction for January 2014, observed visually in the field, ranged between 8 and 35 for “Yellow green” and between 17 and 42 for “white green”. This range was similar for functional classes mapped from the February 2013 data, although mean compaction was slightly higher for February 2013. Giving a recommendation of what flowering density is required for airborne hyperspectral detection of flowering response is not possible given that flowering density is more affected by plant specific canopy background response, phenology and time of observation in the field and from remote sensing as well as color of flowering. In hyperspectral data, leaf level biochemistry variations due to leaf structure and leaf biochemistry differences most profoundly affect canopy level reflectance (Usini et al., 2004). However, phenological and physiological differences are often considered to be more important for canopy level spectral responses (Asner & Martin, 2008; Tits et al., 2012).

In this study, a CVA-based spectral change detection approach was used on pixels that are depicting canopy averages of spectral EM abundances. It was thus assumed that the inaccuracies mentioned are easily managed through the use of bi-temporal hyperspectral data for specific target detection and by using an integrative spectral mapping approach (Buenemann et al., 2011).

High resolution AISA/Eagle hyperspectral airborne imagery is used as input data in this study. Airborne hyperspectral data is often considered better for quantitative scientific analysis than spectral data from Unmanned Aerial Systems (UAS) (Berni, Zarco-Tejada, Suárez, & Fereres, 2009; Hruska, Mitchell, Anderson, & Glenn, 2012). This is because airborne imagery is inherently captured over larger areas with better possibilities for radiometric calibration and mosaicking (Hruska et al., 2012). Using the digital camera imagery for flower mapping, captured in this study simultaneously to the AISA/Eagle imagery, would not be feasible. Camera mounted digital imagery does not have physical reflectance values covering the full electromagnetic spectrum needed for spectral mixture modeling (i.e. linear unmixing) and spectral change detection (i.e. CVA). However, future tests should be done to explore the utility of low cost digital imagery from UAS to map larger flowering trees within smaller study site, using visual interpretation.

The use of hyperspectral data sets with more spectral resolution than the 64 band AISA/Eagle data could have led to less signal-to-noise ratios (SNR) while also exacerbating spectral confusion between landscape features. Spectral confusion is compounded in semi-arid biomes where bi-directional scattering between sub-pixel elements can be very high (Petropoulos, Manevski, & Carlson, 2014).

A higher pixel size could have improved (increased) the SNR for smaller features but this would have also increased per pixel spectral variability with more EM combinations that would have to be accounted for (Plaza et al., 2009; Schaepman et al., 2009). Flowering within plants and tree canopies at the study site also mostly occurred at spatial resolutions smaller than the 60 cm pixel size of the AISA/Eagle data. The use of a straightforward 3 EM sub-pixel mapping method for specific target detection, such as linear spectral unmixing, can thus be deemed appropriately suited to detect sub-pixel flowering effects at the study site using the 60 cm pixel resolution data.

If integrative and accurate flowering activity monitoring methods can be developed using observations from very high resolution spaceborne sensors, then floral calendars can be effectively derived for key sites to guide bee keepers on seasonal floral changes and to help ascertain ‘what to expect from an apiary’ in terms of honey flow. Moreover, if floral cycle and abundance maps are linked to bee foraging preferences, then landscape-based pollination effects can quantified with better accuracies (de Groot, Alkemade, Braat, Hein, & Willemen, 2010).

### Table 2

Overall accuracies for functional flowering plant groups shown as number of pixels and percentages for the January 2014 and February 2013 AISA/Eagle imaging results respectively. Total (n) refers to the total number of reference pixels used for each flowering group in the accuracy assessment.

<table>
<thead>
<tr>
<th>Functional group</th>
<th>Correctly mapped (pixels)</th>
<th>Total (n)</th>
<th>Correctly mapped (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>January 2014 AISA/Eagle data</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crops</td>
<td>81</td>
<td>84</td>
<td>96.4</td>
</tr>
<tr>
<td>White acacia</td>
<td>78</td>
<td>151</td>
<td>51.7</td>
</tr>
<tr>
<td>White forbs</td>
<td>72</td>
<td>73</td>
<td>98.6</td>
</tr>
<tr>
<td>White green</td>
<td>276</td>
<td>315</td>
<td>87.6</td>
</tr>
<tr>
<td>Yellow green</td>
<td>100</td>
<td>129</td>
<td>77.5</td>
</tr>
<tr>
<td>Total pixels</td>
<td>607</td>
<td>752</td>
<td>80.7</td>
</tr>
<tr>
<td>Percentage (All)</td>
<td>80.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>February 2013 AISA/Eagle data</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White acacia</td>
<td>181</td>
<td>211</td>
<td>85.8</td>
</tr>
<tr>
<td>White forbs</td>
<td>75</td>
<td>77</td>
<td>97.4</td>
</tr>
<tr>
<td>White green</td>
<td>137</td>
<td>178</td>
<td>77.0</td>
</tr>
<tr>
<td>Yellow green</td>
<td>32</td>
<td>46</td>
<td>69.6</td>
</tr>
<tr>
<td>Total pixels</td>
<td>425</td>
<td>512</td>
<td>83.0</td>
</tr>
<tr>
<td>Percentage (All)</td>
<td>83</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 3

Producer’s and user’s accuracies (percentages) are shown for all functional flowering plant classes (**All**) mapped using the January 2014 and February 2013 AISA/Eagle data, respectively. Total (n) refers to the total number of reference pixels used for the accuracy assessment.

<table>
<thead>
<tr>
<th>Class (All)</th>
<th>Ground truth data</th>
<th>January 2014</th>
<th>Raw total</th>
<th>User accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>February 2013</td>
<td>425</td>
<td>145</td>
<td>570</td>
<td>74.5</td>
</tr>
<tr>
<td>January 2014</td>
<td>87</td>
<td>607</td>
<td>694</td>
<td>87.4</td>
</tr>
<tr>
<td>Total (n)</td>
<td>512</td>
<td>752</td>
<td>1264</td>
<td></td>
</tr>
<tr>
<td>Producer accuracy</td>
<td>83.0</td>
<td>80.7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Fig. 6.** Box plot distributions for all functional flowering groups together (**All**). Asterisks represented individual data points.
differences in mapping accuracies between functional flowering groups are largely due to flowering compaction as well as background reflectance mixing and intra species phenomenology effects. “White forbs” were most accurately mapped in both ALSA/Eagle data sets, while sparsely flowering plant groups with ‘green’ foliage backgrounds had lower mapping accuracies (i.e. “Yellow green”). The Duncan’s test, as an inter group (class) variance measure to test class similarity, provided an additional insight into the mapping results. Flowering “Crops” in January 2014, for instance, had relatively high overall accuracies but the Duncan’s test revealed that this class was easily confused with “White green”. There is a need to look further into the spatial and temporal variability of the flowering signal or EM in order to investigate upscaling options for the monitoring of the floral cycle using high resolution multispectral data. This would be an immense advantage over field assessment of the floral cycle and be of utmost importance to support bee diversity, hive integrity and decision making around bee hive productivity thus contributing to understanding food security issues in rural Africa.

Acknowledgments

We are grateful to the European Union (DCI-FOOD-2011/023-520) for funding of the project “African reference laboratory (with satellite stations) for the management of pollinator bee diseases and pests for food security” in which context this work was made possible. We sincerely thank Macharia Kabiro for his efforts in field data collections and pre-processing, the University of Helsinki, Department of Geoscience and Geography, for facilitating the hyperspectral data acquisitions and data processing as well as University of Helsinki funding research community GIMMEC–Geoinformatics for monitoring and modeling of environmental change. We also thank Johnston Mutemi and Michael Mulonzya in helping us with the plant species interpretation in the field. We are very thankful to Dolorosa Osogo and other colleagues in icipe for proof reading and technical comments.

References


