A bi-seasonal classification of woody plant species using Sentinel-2A and SPOT-6 in a localised species-rich savanna environment

Emmanuel Fundisia, Solomon G. Tesfamichaela and Fethi Ahmedb

aDepartment of Geography, Environmental Management and Energy Studies, University of Johannesburg, Johannesburg, South Africa; bGeography, Archaeology and Environmental Studies, University of Witwatersrand, Johannesburg, South Africa

ABSTRACT
Sustainable management of biodiversity benefit from cost-effective multi-temporal classification schemes afforded by remote sensing techniques. This study compared classification accuracies of woody plant species (n = 27) and three coexisting land cover types using dry and wet seasons data. Random Forest (RF), Support Vector Machine (SVM) and Deep Neural Network (DNN), were applied to Sentinel-2A and SPOT-6 images. The results showed higher overall classification accuracies for wet season data (65%–72%) for both images and classifiers (DNN, RF and SVM), compared to dry season classification (52%–59%). Near infrared region bands, available in both Sentinel-2A and SPOT-6 imagery, produced high performance for both wet (83%) and dry (80%) seasons. Overall, the findings show potential of multispectral remote-sensing for woody plant species diversity in different seasons. Such a study should be extended to higher frequency species diversity classification, to capture dynamics that may manifest at short time intervals of the year.

ARTICLE HISTORY
Received 2 March 2021
Accepted 24 May 2021

KEYWORDS
Sentinel-2A; SPOT-6; wet and dry seasons; Savanna Woody Plant species; Non-parametric Classifiers

1. Introduction
The savanna ecosystem is considered a rich source of natural resources for firewood, building materials and wildlife habitat, etc. (Mutanga et al. 2016; Axelsson and Hanan 2017; Sosef et al. 2017; Venter et al. 2018; Ondier et al. 2020). Nevertheless, over-exploitation of these resources coupled with climate change is having a devastating effect on the continued integrity of the biodiversity in the savanna region (Marchant 2010; Leeuwis et al. 2018). It is important to have cost-effective and efficient methods for monitoring the savanna ecosystem during dry and wet seasons, to maintain and manage the biodiversity of such an ecosystem. Monitoring biodiversity assists in determining the extent of changes to species composition, such as in their abundance and diversity; it also informs of an incursion of threats, and changes in the condition of vegetation communities.
Remote sensing is one effective technology that has emerged as an alternative to expensive field-based surveys, considering that seasonal analysis would require more than one field survey (Turner et al. 2003; Maxwell et al. 2018).

Optical remote sensing has been used to assess species diversity in different environments, such as in tropical rainforests (Semboli et al. 2014; Wallis et al. 2017), Mediterranean climates (Roberts et al. 2015; Galidaki et al. 2017) and savanna ecosystems (Campos et al. 2018; Silveira et al. 2019). Generally, high accuracy levels are reported in discriminating plant species during the wet season (e.g. Allan et al. 2011; Gessner et al. 2013; Leeuwis et al. 2018; Shaharum et al. 2018; Silveira et al. 2019), due chiefly to the abundance of photosynthetically active species that exhibit a great deal of variability in reflectance during that season. For instance, Ivanova et al. (2019) characterised mixed forests comprising of herbaceous, coniferous and birch-larch forest by exploiting Moderate Resolution Imaging Spectroradiometer (MODIS) data with the advantage of high temporal frequency (< 10-day cycle) that can minimize the effect of cloud contamination during the wet season. Their study reported high accuracy levels; however, there was significant variation in forest types with respect to leaf structures. It is also important to note that MODIS has a coarse spatial resolution that limits its capacity to discriminate species-level diversity. Belay et al. (2013) used Landsat that has better spatial resolution than MODIS imagery, for classifying coniferous and broadleaved trees at the stand level during a wet season. Although they used a better resolution image, they still reported challenges in the effective separation of certain species due to homogeneous spectral characteristics of target species.

Widely used machine learning algorithms such as Random Forest (Breiman 2001) Decision Trees (Quinlan 1986), k-Nearest Neighbors (Fukunaga and Narendra 1975), Gradient Boosting (Friedman 2001) and Support Vector Machines (Cortes and Vapnik 1995) have proved useful by improving accuracies compared to traditional algorithms in woody plant species diversity estimations (Pu and Cheng 2015; Brandt et al. 2016; McKenna et al. 2017; Macintyre et al. 2020). For instance, Jombo et al. (2020) reported high overall accuracies of 81% (SVM) and 84% (RF) for classifying five plant species and three land cover types in a savanna environment, showing a marginally better performance using the latter algorithm. A study by Tesfamichael et al. (2018) showed a slightly better performance of SVM than Gradient Boosting Modeling (GBM) (97 vs. 94%) for discriminating grassland and four shrubby invasive species exhibiting morphologically similar characteristics in a savanna region. Odindi et al. (2016) utilised RF to map the distribution and composition of two alien and indigenous woody vegetation in a savanna environment and reported overall accuracy of 86%. Kganyago et al. (2018) discriminated one species type from three coexisting land cover types in a savanna region using SVM and reported overall accuracy of 83%. Forkuor et al. (2018) assessed the performance of Stochastic Gradient Boosting (SGB), SVM and RF in separating two woody plant species and three land cover types in a savanna region and reported accuracies ranging between 88% and 94%, with SGB performing better than the other classifiers. Such studies have discriminated few woody plant species, and therefore it is important extend the benefits of machine learning algorithms to discerning several species common in complex vegetation environments over different seasons.

A number of studies have exploited optical remote sensing in the absence of atmospheric interference for species diversity classification in dry seasons (Motohka et al. 2010; Cho et al. 2015; Xiao et al. 2015; Möckel et al. 2016; Shaharum et al. 2018; Silveira et al. 2019; Ritz et al. 2020). A common problem with optical remote sensing application in the
dry season relates to challenges in effective discrimination of plant species devoid of chlorophyll content and plant species shedding off their leaves due to absence of moisture in winter (Brandt et al. 2016). This results in a loss of spectral separability between species due to prevalent reflectance from leaf-off branches as well as from background features (Chamaille-Jammes et al. 2006; Cleland et al. 2012; Rocchini et al. 2018; Zimbres et al. 2020). Given this weakness of optical remote sensing in the dry season, careful analysis of the spectral heterogeneity related to species leaf traits and plant phenotyping can be strengthened. Adopting several narrow and contiguous bands can allow the detection of subtle differences in plant species (Guerschman et al. 2009; Fricker et al. 2019; Gao et al. 2019). For instance, Peña et al. (2013) successfully discriminated deciduous and evergreen plants using HySpex-VNIR 1600 airborne-based hyperspectral images. However, the plants had relatively distinct characteristics, with A. caven a deciduous tree measuring approximately 6 m in height and possessing a crown diameter of 0.45 m, and evergreen trees (L. caustica, C. alba and Q. saponaria) measuring approximately 12–30 m in height and having crown diameters of 0.5–1 m. Furthermore, the hyperspectral remote sensing system used in the study is capable of accurately discriminating the vegetation types. Unfortunately, this system is largely commercial and thus unavailable for low-cost monitoring efforts. Generally, it should also be noted that dry period analysis of plant species focuses on a certain developmental stage of plants often devoid of leaves. Such a focus on dry season alone, therefore, limits the importance of optical remote sensing in tracking biodiversity throughout the year. Therefore, there is a need to extend the application of optical remote sensing in plant species diversity classification to both wet and dry seasons.

Monitoring plant species diversity in different seasons provides a temporally sound assessment of biodiversity and ecosystem functioning of a given vegetation environment. In this regard, several studies attempted to classify plant species in different seasons (e.g. Ng et al. 2017; Wendelberger et al. 2018; Gara et al. 2019; Borges et al. 2020; Hunter et al. 2020; Macintyre et al. 2020; Piovan 2020). For instance, Gara et al. (2019) successfully separated broadleaf, conifer, and other landcover types during summer, autumn and winter seasons using Sentinel-2A imagery. Similarly, Chrysafis et al. (2019) reported high accuracy levels in the separation of a mixture of extensive forests, sparse forest, cultivation plants, and pastures, in dry and wet seasons, using Landsat 8. Furthermore, Pu et al. (2018) utilised Pleiades imagery to classify different species in an area dominated by 70% broadleaf evergreen plant species and 30% generic plant species (shrub and turf/lawn/grass) in dry and wet periods. High accuracies reported from the above studies can be attributed to significant differences in vegetation compositions that can be grouped as broadleaf and narrow-leaved, or evergreen and deciduous plants. Foliage dominated by narrow-leaved plants induces spectral responses with little variations, because of similar physiological structure as well as chemical composition of the plant leaves (Bayat et al. 2018). In addition, differentiating such plant species spectrally in different seasons can become complicated in a localised environment where evergreen and deciduous plants are difficult to discern.

There is, therefore, the need to exploit commonly-used multispectral data with improved spatial and spectral characteristics for detecting plants in different seasons. This study aims to investigate the performance of SPOT-6 and Sentinel-2A data for discriminating plant species characterised by similar leaf-level morphological characteristics, in dry and wet seasons. Successful findings from this specific assessment will have great potential in tracking the dynamics of plant species diversity in a localised savanna ecosystem.
2. Materials and methods

2.1. Study area

The Klipriviersberg Nature Reserve south of Johannesburg, South Africa, was used for the study (Figure 1). The reserve was proclaimed for conservation purposes in 1984 and covers approximately 651 hectares. Generally, the vegetation types in the reserve include Andesite Mountain Bushveld and Clay Grassland which are associated with a savanna environment (South African National Biodiversity Institute 2012). The altitude of the area ranges between 1540 m in the south and 1790 m in the north, with a mean altitude of 1653 m, and a coefficient of variation of 3.9%. The mean annual rainfall of the area surrounding the reserve ranges from 624 to 802 mm (Kruger and Nxumalo 2017). The study area experiences warm to hot summers and cold nights in winter, with mean annual temperature ranging between 17°C and 26°C in summer and 5°C and 7°C in winter (MacKellar et al. 2014). The geology types found in the area, which lead to the floristic structure of the reserve, include quartzites, conglomerates, and dolomites (Molezzi et al. 2019).

![Figure 1. Klipriviersberg Nature Reserve and the distribution of sampling plots used in the study. The background image was derived from a WorldView-2 image in false colour composite.](image)

2.2. Field data

In this study, we created a grid of 240 points, distributed at approximately 170 m intervals in the north-south and east-west directions, which covered the whole study area and was generated using the fishnet tool in ArcGIS (ESRI® ArcGIS 10.6, Redlands, California, USA). The point coverage was exported into a Global Position System (GPS)
(Garmin, GPSMAP® 64, Kansas, USA) and located in the field. Field surveys were done in June 2017, representing the dry season, and in November 2017, representing the wet season (MacKellar et al. 2014). A buffer with a 20 m radius was created around each plot center; this size was specified to accommodate multiple pixels of the imagery used in the study. In each plot, different plant species and coexisting land covers were enumerated, totalling 30 classes. All woody plants with a height of two metres or above were counted in each plot. A total of 27 distinct woody plant species were recorded, with a minimum of one and a maximum of nine woody plant species in each plot.

2.3. Remote sensing data and pre-processing

SPOT-6 and Sentinel-2A satellite images were acquired for dry (June 2017) and wet (November 2017) seasons coinciding with the time of the field surveys. The SPOT-6 image was sourced from the South African National Space Agency (SANSA). The imagery has four multispectral bands including red, blue, green (RGB) and near infrared (NIR) in the 0.455–0.89 μm wavelength range (Figure 2) and a panchromatic band in the 0.45 μm–0.75 μm wavelength range. The spatial resolutions of the multispectral and panchromatic images are 6 m and 1.5 m, respectively. Atmospheric correction was applied to all individual bands of the SPOT-6 imagery, using the Dark Object Subtraction (DOS) method (Chavez 1996) implemented in ENVI 5.3 (©2015 Exelis Visual Information Solution Inc., Boulder, CO). Subsequently, the individual bands of SPOT-6 were pansharpened with the panchromatic band using the Gram-Schmidt algorithm, which maximizes image sharpness and minimizes colour distortions (Laben and Brower 2000). The pansharpened bands were combined to create a composite SPOT-6 multispectral image of 1.5 m spatial resolution; one image was generated for each season (wet and dry).

The Sentinel-2A image was downloaded from the European Space Agency Data Hub (https://scihub.copernicus.eu/dhus/). Sentinel-2A has 13 multispectral bands, covering the visible and infrared regions of the electromagnetic spectrum. We excluded the coastal band (0.43 μm–0.45 μm), water vapour band (0.93 μm–0.95 μm), and cirrus band (1.36 μm–1.39 μm) from the Sentinel-2A image, due to their relative sensitivity to atmospheric interferences (Stynch et al. 2019). The remaining 10 bands (Figure 2) covered a
wavelength range of 0.46–2.28 μm. The first four bands out of 10 included the RGB and NIR, which together covered a wavelength range of 0.46–0.90 μm and measured at 10 m spatial resolution. The first four bands out of 10 included RGB and NIR which covered a wavelength range 0.46–0.90 μm measured at 10 m spatial resolution. The additional six bands were vegetation red edge-1 (VRE1), vegetation red edge-2 (VRE2), vegetation red edge-3 (VRE3), vegetation red edge-4 (VRE4), NIR, shortwave infrared-1 (SWIR1) and shortwave infrared-2 (SWIR2) (together spanning 0.71–2.28 μm) were measured at a 20 m spatial resolution. Atmospheric correction was applied to all individual bands of the Sentinel-2A imagery, using a similar method applied to SPOT-6. Subsequently, all the individual bands were stacked to create a multispectral image.

2.4. Training and classification of the remote sensing data

Training of the 30 different classes was performed on the two composite satellite images (SPOT-6 and Sentinel-2A) separately. It should be noted that as the spatial resolution becomes coarser, individual pixels are less likely to capture small features, resulting in a mixed pixel phenomenon (Lu and Weng 2007; Ma et al. 2015; Zhu et al. 2018). This study adopted the nearest neighbour resampling technique on the Sentinel-2A image, in order to resample the pixels at the resolution of the SPOT-6 image (1.5 m). The nearest neighbour resampling technique was chosen because it does not alter values in the output raster data set and is therefore appropriate for categorical data classification (Cassel and Cassel 2013). A total of \( n = 8011 \) (dry season) and \( n = 8080 \) (wet season) points representing 27 woody plant species as well as grass, shrubs and bareland were digitized inside the 240 plots on the two satellite images separately. Digitising of points was guided by field surveys in which a local cartesian coordinate system was used to locate the classes. Finally, points were split into two sets of which 30% was allocated to training — classification and 70% to evaluate the accuracy of the classification. This amounted to \( n = 2408 \) and \( n = 5603 \) samples respectively for the dry season and \( n = 2441 \) and \( n = 5639 \) for wet season respectively. The spatial distribution of the training samples was taken into consideration when splitting the data into training — classification and evaluation sets. The species, along with the proportions allocated to the training and testing of the classifications, are given in Table 1.

Three machine learning classification algorithms—Random Forest (RF), Support Vector Machine (SVM) and Deep Neural Network (DNN)—were applied to train and classify the two datasets. These classification algorithms were implemented using the caret package for RF and SVM (Kuhn 2018), and the H20 package for DNN (Darren 2016) in R (RStudio Team 2020). A RF classifier is an ensemble machine learning method, which uses bootstrap sampling to build multiple decision tree models (Breiman 2001). It was selected for this study because (i) it can handle large datasets, (ii) it is free from normal distribution assumptions, and (iii) it is robust when treating outliers (Cutler et al. 2007). Internally, the RF uses two-thirds of the data (in-bag) for developing the classification model while the remaining one-third, which is referred to as out-of-bag (OOB) data, is used to evaluate the accuracy of the trained model (Breiman 2001). A RF classifier requires the specification of two parameters to generate a prediction model. These parameters include the number of classification trees (ntree) and the number of predicting variables (mtry) used in each node to grow the trees (Cutler et al. 2012). We used a ten-fold cross-validation analysis, which was repeated 10 times to determine the optimal parameters. The explanatory power of the input variables (multispectral bands) was quantified to rank the importance of each band, for classification accuracy.
The SVM classifies features (reflectance of different bands) by trying to identify the optimal decision (separation) boundary that maximizes the margin between two classes (Cortes and Vapnik 1995). Similar to the RF, the SVM does not require the data to have a normal distribution (Maxwell et al. 2018). In addition, SVM performs well when using high dimensional and complex data (Wang and Lin 2014). A nonlinear SVM that utilizes a radial basis function kernel (Cortes and Vapnik 1995) and customized for R (Karatzoglou et al. 2004) was applied in this study, because it accommodates linear and nonlinear relationships between a class and a predictor. The SVM classifier uses two parameters in order to balance the accuracy and reliability of the classification. These parameters includes the cost factor (C) and gamma (γ). The C factor relates to the penalty (cost) of misclassification error and γ determines the influence of a training sample to capture the complexity in the data (Cortes and Vapnik 1995). C and γ were determined by running a ten-fold cross-validation repeated 10 times, similar to the approach applied for selecting optimal parameters of the RF classifier.

Deep learning, which was introduced recently in the field of remote sensing and is a family of machine learning algorithms structured around neural networks, was also used in this study. Deep learning includes multi-layer perceptrons, restricted Boltzmann machines, stacked autoencoders, deep belief networks, and DNN (Schmidhuber 2015). DNN was selected for this study due to its ability to generate non-linear decision boundary, and to learn complex patterns (Lavine and Blank 2009). DNN with error back-propagation was applied using the H2O package in R language (Darren 2016). It uses

| Species name                  | Code | Leaf structure | Dry season Training-Classification | Validation | Wet season Training-Classification | Validation |
|-------------------------------|------|----------------|-----------------------------------|------------|-----------------------------------|------------|
| Acacia caffra                 | AC   | Narrow-leaved  | 265                               | 637        | 265                               | 637        |
| Acacia dealbata               | AD   | Narrow-leaved  | 48                                | 208        | 48                                | 208        |
| Acacia karro                  | AK   | Narrow-leaved  | 177                               | 252        | 177                               | 252        |
| Afrocanthium mundianum        | AM   | Narrow-leaved  | 120                               | 160        | 120                               | 160        |
| Brachylaena rotundata         | BR   | Narrow-leaved  | 104                               | 172        | 104                               | 172        |
| Celtis africana               | CAf  | Narrow-leaved  | 62                                | 139        | 62                                | 139        |
| Celtis australis              | CAu  | Narrow-leaved  | 62                                | 124        | 62                                | 124        |
| Cordyline                     | CO   | Narrow-leaved  | 135                               | 262        | 135                               | 262        |
| Dispyros natalensis           | DN   | Broad-leaved   | 50                                | 284        | 50                                | 284        |
| Dombeya rotundifolia          | DR   | Broad-leaved   | 72                                | 168        | 72                                | 168        |
| Ehretia rigida                | ER   | Narrow-leaved  | 103                               | 207        | 110                               | 214        |
| Eucaea crispa                 | EC   | Narrow-leaved  | 75                                | 138        | 75                                | 138        |
| Gymnosporia buxifolia         | GB   | Narrow-leaved  | 45                                | 240        | 50                                | 245        |
| Heteromorpha arborescens      | HA   | Narrow-leaved  | 38                                | 155        | 38                                | 155        |
| Kiggelaria africana           | KA   | Narrow-leaved  | 88                                | 176        | 88                                | 176        |
| Melia azedarach               | MA   | Narrow-leaved  | 79                                | 105        | 79                                | 105        |
| Olea europaea subs.africana   | OEA  | Narrow-leaved  | 66                                | 175        | 66                                | 175        |
| Pittosporum viridiflorum      | PV   | Narrow-leaved  | 62                                | 152        | 62                                | 152        |
| Populus x canescens           | PC   | Broad-leaved   | 36                                | 147        | 40                                | 153        |
| Rhus lencia                   | RL   | Narrow-leaved  | 59                                | 141        | 59                                | 141        |
| Salix mucronata               | SM   | Narrow-leaved  | 104                               | 151        | 104                               | 151        |
| Sambucus nigra                | SN   | Narrow-leaved  | 48                                | 168        | 48                                | 168        |
| Searssia lepoldcita           | SL   | Narrow-leaved  | 77                                | 126        | 77                                | 126        |
| Searssia pyroides             | SP   | Narrow-leaved  | 67                                | 157        | 67                                | 157        |
| Tarcchonanthus camphoratus    | TC   | Narrow-leaved  | 71                                | 105        | 71                                | 105        |
| Zanthoxylum capense           | ZC   | Narrow-leaved  | 45                                | 114        | 45                                | 114        |
| Ziziphus mucronate            | ZM   | Broad-leaved   | 70                                | 184        | 80                                | 194        |
| Bareland                      | BL   | No leaf        | 54                                | 279        | 54                                | 279        |
| Grassland                     | GL   | No leaf        | 73                                | 133        | 73                                | 133        |
| Shrubs                        | SH   | Mixed          | 53                                | 144        | 60                                | 152        |
| **TOTAL**                     |      |                | 2408                              | 5603       | 2441                              | 5639       |
multi-layer perceptrons that consist of three or more layers; that is, (i) an input layer
which contains one or more processing elements, (ii) hidden layers which are responsible
for the transformation of data between input and output layers, and (iii) an output layer
which stores the results of the network (Hochreiter 1991; Mousavi et al. 2018; Ma et al.
2019). DNN links multiple functions joined together in hierarchically structured neural
networks, which are deeper than three hidden layers (Schmidhuber 2015). We used an
output activation function (hyperbolic tangent function), which is the most frequently
used activation function in deep learning classification (Lavine and Blank 2009; Ma et al.
2019). The function provides zero-centred outputs and allows model parameters to be
more frequently updated in feed-forward neural networks (Lavine and Blank 2009). This
study optimised activation, hidden layers, neurons per layer and epochs hyperparameters
using a grid search approach (Beysolow 2017). A ten-fold cross-validation that was
repeated 10 times was used to determine the best possible combinations of the
hyperparameters.

2.5. Accuracy assessment

Classification accuracies derived from the remotely-sensed data were assessed on the 70%
validation dataset (n = 5603 and n = 5639 for dry and wet seasons, respectively) of the
woody plant species. An error matrix that uses overall, producer’s, and user’s accuracies
(Congalton and Green 2009) was used in the study. Overall accuracy is the probability
that an individual class will be correctly classified. It is measured by the summation of
the true observations plus predicted observations, which is then divided by the total num-
ber of tested classes. The producer’s accuracy indicates the probability of a reference class
species being classified correctly and it is calculated as the number of true observations of
a class divided by the number of true reference observations of the same class. The user’s
accuracy indicates the probability that classified species on the map represent the same
category on the ground. The user’s accuracy is calculated as the number of true observa-
tions of a class (true positives) divided by the number of predicted observations (sum of
true positives and false positives). A Kappa coefficient (Equation 1) was also used as an
additional statistic to assess the quality of classified imagery (Story and Congalton 1986).
The statistical measure is used to control the instances which might have been correctly
classified by chance.

\[
\text{Kappa coefficient} = \frac{P_{\text{observed}} - P_{\text{chance}}}{1 - P_{\text{chance}}} \\
\]

where \( P_{\text{observed}} \) = observed proportion of agreement and \( P_{\text{chance}} \) = proportion expected
by chance.

Finally, the McNemar’s test (McNemar 1947) was used to compare whether or not
there was a significant difference in dry and wet season classification results. This test was
implemented per algorithm; that is, dry season vs. wet season for DNN, dry season vs.
wet season for RF, dry season vs. wet season for SVM. The test was computed using
Equation 2.

\[
Z = \frac{f_{12} - f_{21}}{\sqrt{f_{12} + f_{21}}} \\
\]

where, the square of z follows a chi-square \( \chi^2 \) distribution with 1 degree of freedom. \( f_{12} \)
represents the misclassified number of samples by a classifier (e.g. DNN) in dry season
but classified correctly by the same classifier (e.g. DNN) in wet season. \( f_{21} \) represents total
number of samples classified correctly by a classifier in dry season and not classified correctly by the same classifier in wet season. The pairwise comparison was implemented for overall accuracies as well as for species-level accuracies.

3. Results

3.1. Classification accuracies of dry and wet seasons using RF, SVM and DNN classifiers

Overall classification accuracies of species and coexisting land covers during dry and wet seasons showed that wet season output achieved the highest overall accuracy of 72% (kappa coefficient = 0.68) using a DNN classifier derived from Sentinel-2A data, followed by a 70% accuracy (kappa coefficient = 0.66) with a DNN classifier derived from SPOT-6 data (Figure 3). In contrast, the lowest overall accuracies were recorded in the dry season, at 52% (kappa coefficient = 0.50) using an SVM classifier, followed by an RF classifier with 54% (kappa coefficient = 0.52), both being derived from the same SPOT-6 image. Further analysis of the two seasons and the three classifiers (RF, SVM and DNN) showed overall better results from the wet season data, with mean accuracy of 68% compared to 55% from dry season, which equates to a difference of 13%. Kappa statistics graph is given in supplementary data (Figure S1). Results from the McNemar’s test revealed that wet and dry season classification using DNN classifier had statistically significant difference for Sentinel-2A ($\chi^2 = 8.45; p = 0.003$) and SPOT-6 ($\chi^2 = 7.34; p = 0.007$). The SVM classifier also resulted in a statistically significant difference between the two seasons using Sentinel-2A ($\chi^2 = 4.48; p = 0.023$) and SPOT-6 images ($\chi^2 = 5.36; p = 0.034$). Such significant difference was also observed between the two seasons using RF for Sentinel-2A ($\chi^2 = 4.15; p = 0.042$) and SPOT-6 ($\chi^2 = 3.97; p = 0.048$). Pairwise comparison of individual species using McNemar’s test between wet and dry seasons is provided in supplementary data Table S1. Using Sentinel-2A image, the test showed significant differences between the two seasons for 22, 21 and 20 species for RF, DNN and SVM classifiers, respectively. For the SPOT-6 based classification, the test showed significant differences between the two seasons for 23, 20 and 8 species for RF, DNN and SVM classifiers, respectively.

![Figure 3. Overall accuracies for dry and wet seasons using RF, SVM and DNN derived from Sentinel-2A and SPOT-6 satellite images.](image-url)
Producer’s and user’s accuracies of individual plant species are shown in Figure 4. The producer’s accuracies ranged between 11% (Heteromorpha arborescens) and 96% (Acacia caffra) for different species across wet and dry seasons using RF, SVM and DNN classifiers derived from Sentinel-2A and SPOT-6 images (Figure 4). A closer look at dry season classification results showed that 13 classes (species and coexisting land covers) had a producer’s accuracy exceeding 70% using the three classifiers and Sentinel-2A imagery (Figure 4a). In contrast, 28 classes (species and coexisting land cover types) had a producer’s accuracy greater than 70% using the three classifiers and Sentinel-2A imagery in the wet season (Figure 4b). Furthermore, six species and coexisting land cover types had a producer’s accuracy > 70% using the three classifiers and the SPOT-6 image in the dry season (Figure 4c). Twenty-three species and coexisting land covers had a producer’s accuracy exceeding 70% using the three classifiers and the SPOT-6 image in the wet season (Figure 4d).

The DNN classifier yielded producer’s accuracies varying between 13% (Dombeya rotundifolia) and 80% (A. caffra) in the dry season, as well as between 42% (Melia azedarach) and 91% (A. caffra) in the wet season, using Sentinel-2A imagery. The SVM classifier recorded producer’s accuracies varying between 14% (H. arborescens) and 94% (M. azedarach) in the dry season, as well as 30% (Populus x canescens) and 91% (Acacia delabata) in wet season, using Sentinel-2A imagery. The RF classifier recorded almost similar producer’s accuracies for dry and wet periods, with the values ranging between 35% (Rhus lencia) and 76% (Cordyline) in the dry season. Wet season values ranged between 35% for R. lencia and 87% for Brachylaena rotundata. In both instances, Sentinel-2A imagery was used.

Focusing on user’s accuracy, the three classifiers and Sentinel-2A imagery yielded accuracies ranging between 39% (H. arborescens) and 73% (Acacia Karro) in the dry season, with seven species and coexisting land covers having accuracies greater than 70%. Although the accuracies using the same classifiers and Sentinel-2A ranged between 36% (Tarcchonanthus camphoratus) and 96% (A. caffra) in the wet season, twenty-two species and coexisting land covers had accuracies greater than 70%. The three classifiers and SPOT-6 yielded accuracies ranging between 31% (grassland) and 87% (A. karro) in dry season, with five species and coexisting land covers having accuracies greater than 70%. Although the accuracies using the same SPOT-6 and classifier ranged between 32% (T. camphoratus) and 96% (A. caffra) in wet season, 18 species and coexisting land cover types had accuracies greater than 70%.

The producer’s accuracy was used to compare the relative performance of classifiers between dry and wet seasons in identifying species, since it indicates the probability of a reference class species being classified correctly. Figure 5a shows that wet season classification was better compared to dry season classification for 17 species using DNN, 18 species using SVM and 17 species using RF classifiers, from Sentinel-2A imagery. There was no relative improvement of producer’s accuracy in the dry season over the wet season using Sentinel-2A; consequently, the graph has not been included. The relative improvement of all three classifiers in producer’s accuracy ranged between 1 to 49% when assessing. In contrast, the wet season classification showed an improvement over the dry season for 21 species using DNN and RF, and for 19 species using SVM, using the SPOT-6 image (Figure 5b). In general, the relative improvement in the producer’s accuracy for wet over dry season ranged between 3 and 66% using SPOT-6 (Figure 5b). Interestingly, classification from SPOT-6 imagery for the dry season showed an improvement over the wet season for eight species when applying DNN, for nine species when utilising SVM, and for seven species when using RF (Figure 5c). The relative
improvement in producer’s accuracy in dry season over the wet season ranged between 1 to 31% for a DNN classifier, using SPOT-6 imagery. Comparatively, the improvement in accuracies in dry season over wet season using the SVM classifier ranged between 3% and 22%, while it varied between 1% and 23% for RF classifier (Figure 5c), when applied to SPOT-6 imagery.
3.2. Comparison of wet and dry season based on confusions

Classification performances were further evaluated based on the level of confusion in the identification of species in different seasons. Detailed confusion matrices using DNN, RF and SVM classification types are given in supplementary Table S2. Figure 6 provides the count of species and land cover types against which a species is confused. Lower numbers of species confusion were recorded during the wet season compared to the dry season, showing the superiority of the wet season classification. A species was confused with 10 to 12 other species or land cover types in the wet season when utilising DNN and RF classifiers applied to the Sentinel-2A image (Figure 6a). Results of SPOT-6 derived classification showed that species were on average confused with 12 to 13 other species using DNN and RF classifiers, respectively. The SVM classifier recorded an average of 11 species being confused with other species when using Sentinel-2A imagery, and 12 species when using SPOT-6 imagery. Overall, results of species confusion from the wet season data showed that B. rotundata A. karro and Pittosporum viridiflorum recorded the least number (three) of confusion. Heteromorpha arborescens recorded the highest number (22) of confusion with other species and co-existing land cover types.

Dry season classification results showed that each species was confused with at least 13 other species using DNN, 15 species using SVM, and 17 species using RF, when applied to the Sentinel-2A imagery (Figure 6b). Further poor performance of the SPOT-6 image presented classification results on average 16 species using DNN, 17 species using SVM, and 22 species using the RF classifier. Similar to wet season output, B. rotundata also recorded the lowest number of confusions with other species (confused with three species), whilst A. caffra recorded the highest number (30) of confusions with other land cover types. Overall, the most confusion was recorded in the dry season, compared to the classification in the wet season. Image comparison also highlighted that most of the confusion was recorded using the SPOT-6 image, compared to Sentinel-2A imagery.

Figure 5. Relative producer’s accuracy of wet over dry seasons in identifying species types. Sentinel-2A and SPOT-6 in wet and dry seasons in the y-axes represent satellite images and time period. Species names represented by the two- or three-letter codes are given in Table 1. There was no improvement in dry over wet season for Sentinel-2A.
Variable importance is calculated by the sum of the relative decrease in error when any variable is removed from the classification process (Kuhn 2008). Variable importance was computed for both Sentinel-2A and SPOT-6 images in both wet and dry seasons using the three classifiers (RF, SVM, DNN), in order to observe a general trend in terms of individual band importance. There was a strong similarity between the dry and wet seasons in the important regions of the electromagnetic spectrum of Sentinel-2A imagery. For instance, VRE, NIR and SWIR contributed the most to the classifications (> 50%) using the three classifiers (Figure 7a). SWIR1 contributions in the dry season ranged between 70% and 81%, compared to contributions of 80%–83% in the wet season. The contributions of SWIR2 in the dry season ranged between 55% and 77%, compared to the wet season when it ranged between 53% and 77%. There were also similarities in the contribution of the NIR band between dry and wet seasons, with accuracies exceeding 70%. The VRE bands showed the least contribution in the dry season (31%–76%) compared to the wet season (55%–77%). Similarities were seen in the low contributions (< 50%) of the RGB bands in both seasons. Focusing on SPOT-6 image band contributions (Figure 7b), a similar trend of high values of NIR band was recorded for both dry and wet seasons, with values ranging between 77% and 88% for the former, and between 78% and 83% for the latter. It should be noted that the green band recorded the lowest contributions (< 36%) for both wet and dry seasons. In contrast, both red and blue scored contributions which ranged between 40% and 66% in both wet and dry seasons (Figure 7b).

We assessed the effect of adding spectral indices on classification accuracy. The indices were limited to the bands that contributed the most in the classifications above, including red and NIR for SPOT-6 and Sentinel-2A as well as SWIR1 and 2 for Sentinel-2A. The addition of the Normalised Difference Vegetation Index (NDVI, Rouse et al. 1974) computed as (NIR–Red/NIR + Red) improved the overall classification accuracy of SPOT-6 by...
only 1.5% in the wet season but not in the dry season. When NDVI was added, the overall classification accuracy of Sentinel-2A improved by 1% in the wet season, however it did not improve the overall accuracy in the dry season. A further addition of a Simple Ratio computed as SWIR1/SWIR2 (Henrich et al. 2009) marginally improved the overall classification accuracy of Sentinel-2A by another 2% in the wet season compared to 1% improvement in the dry season. At individual species-level, the improvement in the producer’s accuracies after adding NDVI to SPOT-6 as well as NDVI and Simple Ratio to Sentinel-2A data did not exceed 5% for most species. Considerable improvements were observed for few species that had quite low accuracies (< 20%) without the additions of the spectral indices.

4. Discussion

This study investigated commonly used multispectral data with improved spatial and spectral characteristics for detecting woody plant species in different seasons (wet and dry). The study area is composed of multiple species, including evergreen and deciduous plants with similar morphological characteristics. These characteristics render effective species diversity estimation challenging. Thus, it motivated a study on analysing performance of improved spatial and spectral resolution imagery. Numerous studies have however quantified seasonal species diversity derived from vegetation indices (e.g. Wardlow et al. 2007; Sridhar et al. 2010; Tillack et al. 2014; Adhikari et al. 2016; Wang et al. 2016; Prajesh et al. 2019; Gerard et al. 2020; Wang et al. 2020). In principle, vegetation indices quantify vegetation intensity, and in turn land areas with photosynthetically active woody plant species during wet seasons are misrepresented as high species diversity hotspots. This narrative can misrepresent and foil effective species diversity applications utilising multispectral imagery throughout the year in wet and dry seasons.
4.1. Performance of multispectral image in wet and dry seasons

The results of our study showed high overall accuracy (72%–52%) and kappa coefficient (0.68–0.50) for the wet season compared to the dry season, for species diversity classification (Figure 3). Furthermore, the classification results between wet and dry season showed statistically significant differences. This observation could generally be attributed to an increase in diurnal temperature and maximum photoperiod leading to increased photosynthetic activity and vegetative growth (Eitel et al. 2019). In addition, wet seasons are characterised by high rainfall which contributes to vegetation vigour. Our results are in agreement with Gara et al. (2019) who utilised the same dataset (Sentinel-2A) for seasonal characterisation of leaf canopy although their study focused on vegetation composition that was dominated by herbaceous plants during the summer, spring and autumn seasons. Better accuracies of our study in wet season were not surprising. This is attributed to high photosynthetic activity during the wet season, resulting in plants exhibiting unique chemical composition characteristics of chlorophyll content within the different plant leaves. Similar observations were reported by Kikuzawa (2003) and Wendelberger et al. (2018) and Wang et al. (2016), who mapped species diversity in Prairie grassland with better results during wet seasons.

Although the difference in classification output was not significant among DNN, RF and SVM classifiers (Figure 3), the results can be attributed to different learning approaches and the constraints inherent in each of the algorithms (Bui et al. 2020). For instance, RF is a tree-based classifier that uses repeated sub-setting which groups the training data into categories of increasing similarity (Breiman 2001). The SVM is also designed to split the training data mathematically, using hyperplanes class assignment due to position in multivariate space (Cortes and Vapnik 1995). These classifiers suggest that there is a level of less species variability that makes classification sub-setting inefficient. Thus, tree-based models and the SVM model, which both need well-structured training data in order to provide more accurate models, appear to have less predictive performance than the neural network-based models (Ma et al. 2019). A key advantage of neural networks using deep learning works is in the design which sends the input (the data of images) through different layers of the network, with each network hierarchically defining specific features of images (Abdi 2020). Thus, DNN was able to successfully separate different species better than RF and SVM machine learning algorithms.

Another notable aspect that could have influenced the overall classification performance is the issue of unbalanced training samples (Table 1). Ordinarily, performance of the remote sensing classification assessment produces better results if classes are of almost similar proportions in order to reduce bias towards the most dominant class (Foody 2004; Congalton and Green 2009). Various studies have balanced training data sample proportions in order to improve spectral separation of landcover types (e.g. Olofsson et al. 2014; Romero et al. 2016; Piironen et al. 2017). Abdi (2020) mapped vegetation diversity using different machine learning classification algorithms (Extreme Gradient Boosting (EGB), DNN, SVM and RF) and balanced the training samples regardless of class composition proportions, in order to improve their performance. However sampling design should capture the true representation of identified species in the area of study, for accurate ecological and conservation efforts, rather than focussing on improving accuracy levels.

The Sentinel-2A image produced better results in discriminating multiple narrow-leaved woody plant species in both wet and dry seasons (e.g. *A. caffra*, *A. karro* and *A. mundianum*). Although SPOT-6 has a better spatial resolution compared to Sentinel-2A, the advantage of the latter is due to the improved spectral contents (10 bands versus 4 bands). Such observations were noted by Lewis et al. (2012) who mapped species diversity
(n = 22) using SPOT-5 and Landsat, which has somewhat comparable spectral properties as Sentinel-2A. Hsieh et al. (2001) also reported that the increase in spatial resolution of satellite imagery could result in increased spectral covariance that diminishes class spectral separability. This could explain the superior performance of Sentinel-2A imagery in our study with a 10 m spatial resolution, compared to SPOT-6 with 1.5 m spatial resolution. It should also be noted that improved spectral characteristics of Sentinel-2A data, particularly the presence of four VRE and two SWIR bands, enables the satellites to sense important biophysical parameters of plant species (Arroyo-Mora et al. 2018; Forkuor et al. 2018; Chrysaﬁs et al. 2019). The greater number of bands in Sentinel-2A than in SPOT-6 accentuates effective discrimination of subtle species reflectance variation (Ghoussein et al. 2019).

4.2. Confusion levels of species in wet and dry seasons

The results of the present study comparing discrimination of species diversity in wet and dry seasons demonstrated that the wet season classification produced less confusion among species than the dry season classification (Figure 6). This is due to high photosynthetic activity that allows for the holistic manifestations of leaf spectral properties in wet conditions. Similarly, better performance of wet season classification results was noted by Chrysaﬁs et al. (2019) who mapped five different species and three coexisting landcover types in a Mediterranean climate. Bayat et al. (2018) postulated that plant responses to light differ based on the lighting environment and season. Such factors can explain the different confusion levels between wet and dry season observed in this study (higher species confusion in dry versus wet season as illustrated in Figure 6a and b). Image comparison also highlighted the better classification performance of the Sentinel-2A image than the SPOT-6 image. This could be attributed to the better spectral properties of the Sentinel-2A image. Specifically, Sentinel-2A imagery contains more spectral bands including vegetation-focused channels, many of which are unavailable with SPOT-6. This also implies that the superior spatial resolution of a SPOT-6 image could not compensate for the lack in spectral qualities.

Although inferior to wet season classification, results of species classification in the dry season were encouraging, with producer accuracies exceeding 70% for both Sentinel-2A and SPOT-6 (Figure 5a and c). Similar good performance of classification using dry season was also reported by Hunter et al. (2020), who classified a fewer number of species (n = 9). However, sub-optimal accuracies in dry season classification could be due to the fact that the season is characterised by deciduous plants shedding most of their leaves, rendering plant canopy detection difficult. Another noteworthy morphological characteristic in this study which could have limited the success of species classification is the similarity in size of the species. It should be noted that > 90% of the species identified in the study are narrow-leaved (Table 1), which limits effective species diversity classification.

4.3. Comparison of images based on band importance

A look at the importance of variables in the classifications shows similarities in high contributions from the NIR band of both SPOT-6 and Sentinel-2A images (Figure 7). Such performance is not surprising given that the emissivity rate of the leaf surface of a fully replenished plant (in the wet season) is generally in the range of 0.96 μm–0.99 μm of the electromagnetic spectrum (Xue and Su 2017). The contributions of VRE bands was fairly promising in the classification of plant species with morphologically similar
characteristics. Munyati (2018) classified different plant species composition in savanna and noted that VRE performed generally well in discriminating subtle differences in plant species, because it is more sensitive to variation in chlorophyll content. Macintyre et al. (2020) reported good performance of Sentinel-2A imagery in detecting flowering phenology season of plant species, and attributed it to the presence of VRE bands. Sentinel-2A SWIR bands are centred at 1.61 μm (band 11) and 2.19 μm (band 12), and thus able to detect variability in water content between different tree species (Lukeš et al. 2013; Macintyre et al. 2020). These properties could also explain the good contributions of SWIR bands. The visible range generally showed a similar pattern of low contributions in classifying different plant species in both dry and wet seasons using both images (Figure 7). This could be attributed to the closed-canopy nature of plant species in the study area. Tucker (1979) demonstrated that the reflectance difference in the visible and NIR regions could be related to various properties of plant density, with NIR performing better than the visible range. Poor performances of the bands found in the visible range was also reported by Abdi (2020), who classified different plant species and coexisting landcover types using Sentinel-2A imagery.

5. Conclusions
This study explored the classification accuracies of narrow-leaved woody plant species in dry and wet seasons using SPOT-6 and Sentinel-2A imagery. Overall, classification of plant species had better accuracies in wet season than in the dry season. This is attributed to high photosynthetic activity (vegetation vigour) during the wet season, compared to the dry season. Though not as good as the wet season classification, the dry season classification still yielded promising results. This potentially makes it useful for operational monitoring purposes, due to the fact that images acquired during dry seasons are less affected by atmospheric interferences than those acquired in the wet seasons. Comparing the performances of images, Sentinel-2A resulted in better classification accuracies than the SPOT-6 image in both wet and dry seasons. This was due to the better spectral characteristics of Sentinel-2A imagery, and despite the better spatial resolution provided by SPOT-6.

Variable importance results showed consistently high contributions of the NIR and longer wavelength regions to the species classifications using both images in dry and wet seasons. Such performances indicate the adequacy of VRE, NIR and SWIR range bands for narrow-leaved plant species classification in both wet and dry seasons. Our study classified species diversity comparing only wet and dry season; however, it is important to expect the existence of species that exhibit dynamics at shorter intervals than the two generic seasons considered in the study. Therefore, classification of species type should be extended to higher temporal frequency analysis in order to capture the dynamics for improved biodiversity monitoring and management.

Acknowledgments
The University of Johannesburg (South Africa) provided the necessary financial and material support to undertake the study. We thank the National Research Foundation (NRF) of South Africa for supporting the first author through student scholarship program (Reference SFH150803134516). The authors also thank Bishop Ngobeli (Manager of the Klipriviersberg Nature Reserve) for unlimited access to the reserve and for additional transportation support to and from the reserve. We are grateful to Tumelo Molaba for all the assistance during the field survey.
Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors confirm that the data supporting the findings of this study are available within the article and its supplementary materials (Tables S1 and S2, Figure S1).

References

Abdi AM. 2020. Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data. GIScience Remote Sens. 57(1):1–20.

Adhikari H, Heiskanen J, Maeda EE, Pelliotta PKE. 2016. The effect of topographic normalization on fractional tree cover mapping in tropical mountains: an assessment based on seasonal Landsat time series. Int J Appl Earth Obs Geoinf. 52:20–31.

Allan E, Weisser W, Weigelt A, Roscher C, Fischer M, Hillebrand H. 2011. More diverse plant communities have higher functioning over time due to turnover in complementary dominant species. Proc Natl Acad Sci USA. 108(41):17034–17039.

Arroyo-Mora JP, Kalacska M, Soffer R, Ifimov G, Leblanc G, Schaal ES, Lucasus O. 2018. Evaluation of phenospectral dynamics with Sentinel-2A using a bottom-up approach in a northern ombrotrophic peatland. Remote Sens Environ. 216:544–560.

Axelsson CR, Hanan NP. 2017. Patterns in woody vegetation structure across African savannas. Biogeosciences. 14(13):3239–3252.

Bayat L, Arab M, Aliniaeifard S, Seif M, Lastochkina O, Li T. 2018. Effects of growth under different light spectra on the subsequent high light tolerance in rose plants. AoB Plants. 10(5):plp052.

Belay TA, Totland T, Moe SR. 2013. Woody vegetation dynamics in the rangelands of lower Omo region, southwestern Ethiopia. J Arid Environ. 89:94–102.

Beysolow TI. 2017. Introduction to deep learning using R: a step-by-step guide to learning and implementing deep learning models using R. San Francisco (California): Academic Press.

Borges J, Higginbottom TP, Symeonakis E, Jones M. 2020. Sentinel-1 and sentinel-2 data for savannah land cover mapping: optimising the combination of sensors and seasons. Remote Sens. 12(23):3862.

Brandt M, Hiernaux P, Tagesson T, Verger A, Rasmussen K, Diouf AA, Mbow C, Mougin E, Fensholt R. 2016. Woody plant cover estimation in drylands from earth observation based seasonal metrics. Remote Sens Environ. 172:28–38.

Breiman L. 2001. Random forests. Mach Learn. 45(1):5–32.

Bui DT, Tsangaratos P, Nguyen VT, Liem NV, Trinh PT. 2020. Comparing the prediction performance of a deep learning neural network model with conventional machine learning models in landslide susceptibility assessment. Catena. 188:104426.

Campos VE, Gatica GM, Cappa FM, Giannoni SM, Campos CM. 2018. Remote sensing data to assess compositional and structural indicators in dry woodland. Ecol Indic. 88:63–70.

Cassel KW, Cassel KW. 2013. Image processing and data analysis. In Variational methods with applications in science and engineering.

Chamaillé-Jammes S, Fritz H, Murindagomo F. 2006. Spatial patterns of the NDVI-rainfall relationship at the seasonal and interannual time scales in an African savanna. Int J Remote Sens. 27(23):5185–5200.

Chavez PS. 1996. Image-based atmospheric corrections – revisited and improved. Photogramm Eng Remote Sens. 62(9):1025–1035.

Cho MA, Malahlela O, Ramoelo A. 2015. Assessing the utility WorldView-2 imagery for tree species mapping in South African tropical humid forest and the conservation implications: Dukuduku forest patch as case study. Int J Appl Earth Obs Geoinf. 38:349–357.

Chrysas I, Mallinis G, Tsakiri M, Patias P. 2019. Evaluation of single-date and multi-seasonal spatial and spectral information of Sentinel-2 imagery to assess growing stock volume of a Mediterranean forest. Int J Appl Earth Obs Geoinf. 77:1–14.

Cleland EE, Allen JM, Crimmins TM, Dunne JA, Pau S, Travers SE, Zavaleta ES, Wolkovich EM. 2012. Phenological tracking enables positive species responses to climate change. Ecology. 93(8):1765–1771.

Congalton RG, Green K. 2009. Assessing the accuracy of remotely sensed data. Principles and practices, 2nd ed. Boca Ratón (FL): CRC Press. Taylor and Francis.
Cortes C, Vapnik V. 1995. Support-vector networks. Mach Learn. 20(3):273–297.
Cutler A, Cutler DR, Stevens JR. 2012. Random forests. In Cutler A, Cutler DR, Stevens JR, editors. Ensemble machine learning: methods and applications. Boston (MA): Springer; p. 157–175.
Cutler DR, Edwards TC, Beard KH, Cutler A, Hess KT, Gibson J, Lawler JJ. 2007. Random forests for classification in ecology. Ecology. 88(11):2783–2792.
Darren C. 2016. Practical machine learning with H2O: powerful, scalable techniques for deep learning and AI. In: N Tache, editor. Practical machine learning with Python, 1st ed. California: O’Reilly Media.
Eitel JUH, Maguire AJ, Boelman N, Vierling LA, Griffin KL, Jensen J, Magney TS, Mahoney PJ, Meddens AJH, Silva C, et al. 2019. Proximal remote sensing of tree physiology at northern treeline: do late-season changes in the photochemical reflectance index (PRI) respond to climate or photoperiod? Remote Sens Environ. 221:340–350.
Foody GM. 2004. Thematic map comparison: evaluating the statistical significance of differences in classification accuracy. Photogramm Eng Remote Sens. 70(5):627–633.
Forkuor G, Dimobe K, Serme I, Tondoh JE. 2018. Landsat-8 vs. Sentinel-2: examining the added value of sentinel-2’s red-edge bands to land-use and land-cover mapping in Burkina Faso. GIScience Remote Sens. 55(3):331–354.
Fricker GA, Ventura JD, Wolf JA, North MP, Davis FW, Franklin J. 2019. A convolutional neural network classifier identifies tree species in mixed-conifer forest from hyperspectral imagery. Remote Sens. 11(19):2326.
Friedman JH. 2001. Greedy function approximation: a gradient boosting machine. Ann Statis. 29:1189–1232.
Fukunaga K, Narendra PM. 1975. A branch and bound algorithm for computing k-nearest neighbors. IEEE Trans Comput. C-24(7):750–753.
Galidaki G, Zianis D, Gitas I, Radoglou K, Karathanassi V, Tsakiri-Strati M, Woodhouse I, Mallinis G. 2017. Vegetation biomass estimation with remote sensing: focus on forest and other wooded land over the Mediterranean ecosystem. Int J Remote Sens. 38(7):1940–1966.
Gao J, Meng B, Liang T, Feng Q, Ge J, Yin J, Wu C, Cui X, Hou M, Liu J, et al. 2019. Modeling alpine grassland forage phosphorus based on hyperspectral remote sensing and a multi-factor machine learning algorithm in the east of Tibetan Plateau, China. ISPRS J Photogramm Remote Sens. 147:104–117.
Gara TW, Darvishzadeh R, Skidmore AK, Wang T, Heurich M. 2019. Accurate modelling of canopy traits from seasonal Sentinel-2 imagery based on the vertical distribution of leaf traits. ISPRS J Photogramm Remote Sens. 157:108–123.
Gerard FF, George CT, Hayman G, Chavana-Bryant C, Weedon GP. 2020. Leaf phenology amplitude derived from MODIS NDVI and EVI: maps of leaf phenology synchrony for Meso- and South America. Geosci Data J. 7(1):13–26.
Gessner U, Machwitz M, Conrad C, Dech S. 2013. Estimating the fractional cover of growth forms and bare surface in savannas. A multi-resolution approach based on regression tree ensembles. Remote Sens Environ. 129:90–102.
Ghoussein Y, Nicolas H, Haury J, Fadel A, Pichelin P, Hamdan HA, Faour G. 2019. Multitemporal remote sensing based on an FVC reference period using sentinel-2 for monitoring Eichhornia crassipes on a Mediterranean river. Remote Sens. 11(16):1856.
Guerschman JP, Hill MJ, Renzullo LJ, Barrett DJ, Marks AS, Botha EJ. 2009. Estimating fractional cover of photosynthetic vegetation, non-photosynthetic vegetation and bare soil in the Australian tropical savanna region upsampling the EO-1 Hyperion and MODIS sensors. Remote Sens Environ. 113(5):928–945.
Henrich V, Jung A, Götze C, Sandow C, Thirkow DG. 2009. Development of an online indices database: motivation, concept and implementation. 6th EARSeL Imaging Spectroscopy.
Hochreiter S. 1991. Untersuchungen zu dynamischen neuronalen Netzen [master’s thesis]. Munchen (Germany): Institut für Informatik, Technische Universität.
Hsieh PF, Lee LC, Chen NY. 2001. Effect of spatial resolution on classification errors of pure and mixed pixels in remote sensing. IEEE Trans Geosci Remote Sens. 39(12):2657–2663.
Hunter FDL, Mitchard ETA, Tyrrell P, Russell S. 2020. Inter-seasonal time series imagery enhances classification accuracy of grazing resource and land degradation maps in a savanna ecosystem. Remote Sens. 12(1):198.
Ivanova Y, Kovalev A, Yakubailik O, Soukhovolsky V. 2019. The use of satellite information (MODIS/Aqua) for phenological and classification analysis of plant communities. Forests. 10(7):561.
Jombo S, Adam E, Byrne MJ, Newete SW. 2020. Evaluating the capability of Worldview-2 imagery for mapping alien tree species in a heterogeneous urban environment. Cogent Soc Sci. 6(1):1754146.
Karatzoglou A, Hornik K, Smola A, Zeileis A. 2004. kernlab – an S4 package for kernel methods in R. J Stat Soft. 11(9):1–20.

Kganyago M, Odindi J, Adjlrorolo C, Mhangara P. 2018. Evaluating the capability of Landsat 8 OLI and SPOT 6 for discriminating invasive alien species in the African Savanna landscape. Int J Appl Earth Obs Geoinf. 67:10–19.

Kikuzawa K. 2003. Phenological and morphological adaptations to the light environment in two woody and two herbaceous plant species. Funct Ecol. 17(1):29–38.

Kruger AC, Nxumalo MP. 2017. Historical rainfall trends in South Africa: 1921–2015. WSA. 43(2): 285–297.

Kuhn M. 2008. Building predictive models in R using the caret package. J Stat Soft. 28(5):1–26.

Kuhn M. 2018. R Package: caret, Version. 6.0-81. CRAN.

Laben C, Brower B. 2000. Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening. United States Patent 6.

Lavine BK, Blank TR. 2009. Feed-forward neural networks. In: Brown S, Tauler R, Walczak B, editors. Comprehensive chemometrics. Amsterdam: Elsevier; p. 571–586.

Leeuwis T, Peel M, De Boer WF. 2018. Complexity in African savannas: direct, indirect, and cascading effects of animal densities, rainfall and vegetation availability. PLoS One. 13(5):e0197149.

Lewis D, Phinn S, Pfitzner K. 2012. Pixel-based image classification to map vegetation communities using SPOT5 and Landsat5 thematic mapper data in a tropical savanna, northern Australia. Can J Remote Sens. 38(5):570–585.

Lu D, Weng Q. 2007. A survey of image classification methods and techniques for improving classification performance. Int J Remote Sens. 28(5):823–870.

Luks P, Stenberg P, Raatiainen M, Möttus M, Vanhatalo KM. 2013. Optical properties of leaves and needles for boreal tree species in Europe. Remote Sens Lett. 4(7):667–676.

Ma L, Liu Y, Ma X. 2015. Training set size, scale, and features in geographic object-based image analysis of very high resolution unmanned aerial vehicle imagery. ISPRS J Photogramm Remote Sens. 102:14–27.

Ma L, Liu Y, Zhang X, Ye Y, Yin G, Johnson BA. 2019. Deep learning in remote sensing applications: a meta-analysis and review. ISPRS J Photogramm Remote Sens. 152:166–177.

Macintyre P, van Niekerk A, Mucina L. 2020. Efficacy of multi-season Sentinel-2 imagery for compositional vegetation classification. Int J Appl Earth Obs Geoinf. 85:101980.

Marchant R. 2010. Understanding complexity in savannas: climate, biodiversity and people. Curr Opin Environ Sustain. 2(1-2):101–108.

Maxwell AE, Warner TA, Fang F. 2018. Implementation of machine-learning classification in remote sensing: an applied review. Int J Remote Sens. 39(9):2784–2817.

McKenna P, Erskine PD, Lechner AM, Phinn S. 2017. Measuring fire severity using UAV imagery in semi-arid central Queensland, Australia. Int J Remote Sens. 38(14):4244–4264.

McNemar Q. 1947. Note on the sampling error of the difference between correlated proportions or percentages. Psychometrika. 12(2):153–157.

Möckel T, Dalmayne J, Schmid BC, Prentice HC, Hall K. 2016. Airborne hyperspectral data predict finescale plant species diversity in grazed dry grasslands. Remote Sens. 8(2):133.

Molezzi MG, Hein KAA, Manzi MSD. 2019. Mesoarchaean-palaeoproterozoic crustal-scale tectonics of the central Witwatersrand basin – interpretation from 2D seismic data and 3D geological modelling. Tectonophysics. 761:65–85.

Motohka T, Nasahara KN, Oguma H, Tsuchida S. 2010. Applicability of Green-Red Vegetation Index for remote sensing of vegetation phenology. Remote Sens. 2(10):2369–2387.

Mousavi SS, Schukat M, Howley E. 2018. Deep reinforcement learning: An overview. Proceedings of SAI Intelligent Systems Conference. p. 1–17.

Munyati C. 2018. Optimising multiresolution segmentation: delineating savannah vegetation boundaries in the Kruger National Park, South Africa, using Sentinel 2 MSI imagery. Int J Remote Sens. 39(18):5997–6019.

Mutanga O, Dube T, Ahmed F. 2016. Progress in remote sensing: vegetation monitoring in South Africa. South Afr Geogr J. 98(3):461–471.

Ng WT, Rima P, Einzmann K, Immitzer M, Atzberger C, Eckert S. 2017. Assessing the potential of sentinel-2 and pleiades data for the detection of Prosopis and Vachellia spp. in Kenya. Remote Sens. 9(1):74.

Odindi J, Mutanga O, Rouget M, Hlanguza N. 2016. Mapping alien and indigenous vegetation in the KwaZulu-Natal Sandstone Sourveld using remotely sensed data. Bothalia. 46(2):1–9.
Olofsson P, Foody GM, Herold M, Stehman SV, Woodcock CE, Wulder MA. 2014. Good practices for estimating area and assessing accuracy of land change. Remote Sens Environ. 148:42–57.

Ondier J, Okach DO, John OC, Otieno DO. 2020. Influence of rainfall amount and livestock grazing on soil respiration in a moist Kenyan savannah. Afr J Ecol. 58(1):92–99.

Peña MA, Cruz P, Roig M. 2013. The effect of spectral and spatial degradation of hyperspectral imagery for the Sclerophyll tree species classification. Int J Remote Sens. 34(20):7113–7130.

Piironen R, Heiskanen J, Maeda E, Viinikka A, Pelliikka P. 2017. Classification of tree species in a diverse African Agroforestry landscape using imaging spectroscopy and laser scanning. Remote Sens. 9(9):875.

Piovan SE. 2020. Remote sensing. In Springer geography. Cham (Switzerland): Springer. The geohistorical approach; Springer Geography. Cham (Switzerland): Springer: 2020; p. 171–197.

Prajesh PJ, Kannan B, Pazhanivelan S, Ragunath KP. 2019. Monitoring and mapping of seasonal vegetation trend in Tamil Nadu using NDVI and NDWI imagery. JANS. 11(1):54–61.

Pu R, Cheng J. 2015. Mapping forest leaf area index using reflectance and textural information derived from WorldView-2 imagery in a mixed natural forest area in Florida, US. Int J Appl Earth Obs Geoinf. 42:11–23.

Pu R, Landry S, Yu Q. 2018. Assessing the potential of multi-seasonal high resolution Pléiades satellite imagery for mapping urban tree species. Int J Appl Earth Obs Geoinf. 71:144–158.

Quinlan JR. 1986. Induction of decision trees. Mach Learn. 1(1):81–106.

RStudio Team. 2020. RStudio: Integrated Development for R. RStudio, PBC, Boston, MA. http://www.rstudio.com

Ritz E, Bjerke JW, Tømmervik H. 2020. Monitoring winter stress vulnerability of high-latitude understory vegetation using intraspecific trait variability and remote sensing approaches. Sensors. 20(7):2102.

Roberts DA, Dennison PE, Roth KL, Dudley K, Hulley G. 2015. Relationships between dominant plant species, fractional cover and land surface temperature in a Mediterranean ecosystem. Remote Sens Environ. 167:152–167.

Rocchini D, Luque S, Pettorelli N, Bastin L, Doktor D, Faedi N, Fehlauer H, Fêret JB, Foody GM, Gavish Y, et al. 2018. Measuring β-diversity by remote sensing: a challenge for biodiversity monitoring. Methods Ecol Evol. 9(8):1787–1798.

Romero A, Gatta C, Camps-Valls G. 2016. Unsupervised deep feature extraction for remote sensing image classification. IEEE Trans Geosci Remote Sens. 54(3):1349–1362.

Rouse JW, Haas RH, Schell JA, Deering DW. 1974. Monitoring vegetation systems in the Great Plains with ERTS. Third ERTS Symposium. Washington (DC): NASA.

Schmidhuber J. 2015. Deep Learning in neural networks: an overview. Neural Netw. 61:85–117.

Semboli O, Beina D, Closset-Kopp D, Gourlet-Fleury S, Decocq G. 2014. Does long-term monitoring of tropical forests using permanent plots provide unbiased results? Appl Veg Sci. 17(4):737–743.

Shaharum NSN, Shafri HZM, Gambo J, Abidin FAZ. 2018. Mapping of Krau Wildlife Reserve (KWR) protected area using Landsat 8 and supervised classification algorithms. Remote Sens Appl Soc Environ. 10:24–35.

Silveira EM, de O, Terra M, de CNS, ter Steege H, Maeda EE, Acrebi Júnior FW, Scolforo JRS. 2019. Carbon-diversity hotspots and their owners in Brazilian southeastern Savanna, Atlantic Forest and Semi-Arid Woodland domains. For Ecol Manage. 452:117575.

Sosef MSM, Dauby G, Blach-Overgaard A, van der Burgt X, Catarino L, Damen T, Deblauwe V, Dessein S, Dransfield J, Droissart V, et al. 2017. Exploring the floristic diversity of tropical Africa. BMC Biol. 15(1):15.

South African National Biodiversity Institute. 2012. 2012 vegetation map of South Africa, Lesotho and Swaziland [Vector geospatial dataset] 2. Pretoria: Biodiversity GIS Website.

Sridhar BBM, Vincent RK, Clapham WB, Sritharan SI, Osterberg J, Neale CMU, Watts DR. 2010. Mapping saltcedar (Tamarix ramosissima) and other riparian and agricultural vegetation in the lower Colorado River region using multi-spectral landsat tm imagery. Geocarto Int. 25(8):649–662.

Story M, Congalton RG. 1986. Remote sensing brief accuracy assessment: a user’s perspective. Photogramm Eng Remote Sens. 52(3):397–399.
Tucker CJ. 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens Environ. 8(2):127–150.

Turner W, Spector S, Gardiner N, Fladeland M, Sterling E, Steininger M. 2003. Remote sensing for biodiversity science and conservation. Trends Ecol Evol. 18(6):306–314.

Venter ZS, Cramer MD, Hawkins HJ. 2018. Drivers of woody plant encroachment over Africa. Nat Commun. 9(1):1–7.

Wallis CIB, Brehm G, Donoso DA, Fiedler K, Homeier J, Paulsch D, Süßenbach D, Tiede Y, Brandl R, Farwig N, et al. 2017. Remote sensing improves prediction of tropical montane species diversity but performance differs among taxa. Ecol Indic. 83:538–549.

Wang PW, Lin CJ. 2014. Iteration complexity of feasible descent methods for convex optimization. J Mach Learn Res. 15(2014):1523–1548.

Wang R, Gamon JA, Emmerton CA, Springer KR, Yu R, Hmimina G. 2020. Detecting intra- and interannual variability in gross primary productivity of a North American grassland using MODIS MAIAC data. Agric For Meteorol. 281:107859.

Wang R, Gamon JA, Montgomery RA, Townsend PA, Zygielbaum AI, Bitan K, Tilman D, Cavender-Bares J. 2016. Seasonal variation in the NDVI-species richness relationship in a prairie grassland experiment (Cedar Creek). Remote Sens. 8(2):128.

Wardlow BD, Egbert SL, Kastens JH. 2007. Analysis of time-series MODIS 250 m vegetation index data for crop classification in the U.S. Central Great Plains. Remote Sens Environ. 108(3):290–310.

Wendelberger KS, Gann D, Richards JH. 2018. Using bi-seasonal worldview-2 multi-spectral data and supervised random forest classification to map coastal plant communities in everglades national park. Sensors. 18(3):829.

Xiao L, Feng M, Yang W, Ding G. 2015. Estimation of water content in winter wheat (Triticum aestivum L.) and soil based on remote sensing data–vegetation index. Commun Soil Sci Plant Anal. 46(14):1827–1839.

Xue J, Su B. 2017. Significant remote sensing vegetation indices: a review of developments and applications. J Sens. 2017:1–17.

Zhu L, Suomalainen J, Liu J, Hyypää J, Kaartinen H, Haggren H. 2018. A review: remote sensing sensors. In: Rustamov RB, editor. London: IntechOpen Multi-purposeful application of geospatial data; 19–42.

Zimbres B, Shimbo J, Bustamante M, Levick S, Miranda S, Roitman I, Silvério D, Gomes L, Fagg C, Alencar A. 2020. Savanna vegetation structure in the Brazilian Cerrado allows for the accurate estimation of aboveground biomass using terrestrial laser scanning. For Ecol Manage. 458:117798.