Towards a semi-automated mapping of Australia native invasive alien *Acacia* trees using Sentinel-2 and radiative transfer models in South Africa

Cecilia Masemola[a,b],*, Moses Azong Cho[d,c] and Abel Ramoelo[e,f]

[a]Discipline of Geography, School of Agricultural, Earth and Environmental Sciences, University of KwaZulu-Natal, Private Bag X01, Scottsville, Pietermaritzburg 3209, South Africa
[b]College of Agriculture and Environmental Science, University of South Africa, Pretoria, South Africa
[c]Earth Observation Research Group, Natural Resources and Environment, Council for Scientific and Industrial Research, Pretoria, South Africa
[d]Forest Science Postgraduate Programme, Department of Plant and Soil Science, University of Pretoria, South Africa
[e]Conservation Services, South African National Parks (SANParks), P.O. Box 787, Pretoria 0001, South Africa
[f]Vulnerability Assessment Centre, University of Limpopo, Sovenga, South Africa.

*Corresponding author. Masemolac@ukzn.ac.za

Abstract

Invasive alien plants (IAPs) threaten biodiversity and critical ecosystem services worldwide. There is, therefore, an urgent need to develop intervention measures to control the spread of IAPs. Efforts to control and monitor the spread of IAPs would require their current and detailed distribution over a large geographic area. Recently launched multispectral instrument on-board Sentinel-2 provides free data with good spatiotemporal and spectral resolution, compared to Landsat datasets. The Sentinel-2 dataset, therefore, can be a useful source of the IAPs spatial information required for detection and monitoring purposes. We combined Sentinel-2 data with a radiative transfer model to discriminate IAPs (*Acacia mearnsii* and *Acacia dealbata*) from surrounding native tree species in Van Reenen, KwaZulu-Natal, South Africa. The forward mode of combined PROSPECT leaf optical properties model and SAIL canopy bidirectional reflectance model, also referred to as PROSAIL was used to simulate reflectance corresponding to bands of Sentinel-MSI, while the PROSAIL model inversion retrieved leaf area index (LAI) and canopy chlorophyll contents (CCC) of the IAPs and native species. Both reflectance and retrieved properties were used to map the distribution of the species within the study area. Our results showed that *A. mearnsii* and *A. dealbata* could be accurately discriminated from the surrounding native trees using integrated PROSAIL Sentinel-2 based model. We found that CCC– and LAI-based (% accuracy = 92.8%, 91.4% for CCC and LAI, respectively) modelling produced a higher classification accuracy than field sampling-based modelling (Accuracy = 90.2% (IAP), 82.2% (NAT) and kappa coefficient = 0.84 (IAP), 0.78 (NAT)). Simulated bands corresponding to Sentinel-2 data, on the other hand, produced species maps comparable to field sampling-based maps. Overall, the integrated PROSAIL Sentinel-2 inversion approach proved suitable for detecting and mapping IAPs over a large area. Due to the high spatiotemporal coverage of Sentinel-2, satellite images, the model developed showed the potential to contribute to the IAPs monitoring systems.

Keywords: Invasive alien plant; Radiative Transfer Model; PROSAIL; Sentinel-2; Leaf Area Index; Canopy Chlorophyll Content

1. Introduction

The world is experiencing a loss of biodiversity because of invasive alien plants (IAPs) (Simberloff et al., 2013). IAPs are destructive because of their ability to compete with native vegetation for soil nutrients, water and light. It has been argued that the competitiveness of IAPs is associated with their unique biological traits (Morris et al., 2011, Große-Stoltenberg et al., 2018). For example, Australian native *Acacia spp* (Maitre et al., 2011), impact native species by discharging large amounts of nitrogen-rich litter (Makkonen et al., 2012), thus, changes nutrient chemistry in such a way that it become toxic to native species (Morris et al., 2011, Rascher et al., 2012, Hellmann et al., 2017). Furthermore, the species have also been characterised as prolific water, and resource users (Chamier et al., 2012, van Wilgen and Wilson, 2018) and they can alter carbon cycles (Rascher et al., 2012). These characteristics enable the Australian *Acacia spp*. to adapt and establish in non-native habitats. The distinctive characteristics of Australian *Acacias* have enabled them to be spectrally distinguished from non-invasive species using air- or space-borne remote sensing data (Pandey, Tate and Balzter, 2014).

In remote sensing, traits variations between species are used for their classification (Jiménez and Díaz-Delgado, 2015). Currently, the most established and frequently used methods to map species traits are empirical statistical regression approaches (Großkinsky et al., 2015, Chance et al., 2016, Große-Stoltenberg et al., 2016, Skowronek et al., 2017, Große-Stoltenberg et al., 2018). Although satisfactory application of statistical models has been widely demonstrated, the empirical relationships between plant traits and reflectance data are site-and species-specific (Punalekar et al. 2018). In addition, the techniques are dependent on site-specific calibrations (Punalekar et al. 2018) and require a considerable ground-based species database to validate every new situation. Hence, empirical models lack robust portability and transferability. Furthermore, they are highly accurate when used on hyperspectral data (Thenkabail, 2015, Thenkabail and Lyon, 2016), and these data are costly for frequent mapping of IAPs over a larger area (Thenkabail et al., 2019). Recently, efforts have been undertaken to integrate physically based radiative transfer models (RTMs) inversion techniques for mapping plant traits and species distribution (Verrelst et al., 2015a, Verrelst et al., 2015b).

Radiative transfer models adopt the same principles as that of radiative transfer (RT) processes that take place during the radiation propagation within vegetation canopies (Goel, 1988, Baret and Guyot, 1991; Dorrigo *et al.* 2008; Rivera et al., 2014). The models simulate canopy spectral reflectance based on the pre-identified vegetation architecture (LAI, Leaf Angle), leaf biochemical parameters (Chlorophyll, Leaf structure), background soil structure (Soil Brightness, Roughness) and observation geometry of the plant and satellite (Goel et al., 1988). There are two main categories of RTMs, namely, homogeneous and heterogeneous models also known as 1-Dimensional and 3-Dimensional RTMs, respectively. The former is for landscape that is represented by a constant horizontal distribution of absorbing and scattering elements (Ross 1989). While the latter is used for non-uniform distribution of features of the landscape (Verrelst et al. 2012). Among all the RTMs, the Scattering by Arbitrary Inclined Leaves (SAIL) canopy bidirectional reflectance model (Verhoef 1984) and the PROSPECT leaf optical properties model (Jacquemoud and Baret 1990) are the most popular homogenous (1-Dimensional) models. PROSPECT model is the 1-

D RTM that describe the optical properties of plant leaves from the visible (400 nm) to the shortwave infrared (2500 nm). Unlike empirical models, RTMs are not site and sensor-specific; as a result, they are robust and transferable (Baret and Guyot, 1991, Dorigo, 2008, Rivera-Caicedo et al., 2017).

Until recently, RTMs have been mainly used to quantify physical and biochemical properties of vegetation (Verrelst et al., 2013, Lázaro-Gredilla et al., 2014, Van Wittenberghe et al., 2014, Delegido et al., 2015, Verrelst et al., 2015a, Verrelst et al., 2015b, Rivera-Caicedo et al., 2017). The study by (Goel, 1988) was the first to invert SAIL to retrieve canopy architecture (Leaf Area Index (LAI), leaf angle distribu- tion) on soybean. Recently, Féret et al. (2011) optimised vegetation spectral indices and retrieval methods for quantifying leaf properties using RTMs. Only a few studies have demonstrated interest in developing RTMs based operational algorithms for species discrimination at a large scale (Féret and Asner, 2011, Verrelst et al., 2012; Rivera et al., 2013; Sun et al. 2018). Sun et al. 2018 retrieved leaf biochemical properties beased on PROSPECT model inversion and different spectral information using LOPEX and ANGERS experimental data. Verrelst et al., 2015a, Verrelst et al., 2015b indicated the capability of Sentinel-2 for LAI retrieval using physical retrieval methods. Recently Xu et al. (2019) successfully inverted accurate estimates of rice canopy chlorophyll contents and LAI with coupling of radiative transfer and Bayesian network models in Rugao, China.

Relevent to this study, is Verrelst et al. (2012) study that quantified vegetation structure in a complex forest ecosystem based on LAI retrieved using combined 1D and 3D RT model and land cover map produced from CHRIS/PROBA multi-angular observations. Verrelst et al. (2012) demonstrated improvements afforded by the three-dimensional RT models compared to one-dimensional (1D)-Flight RTM version for species distribution at landscape scale. Although 3D-RT models are adequate for heterogeneous ecosystems requires a considerable number of parameters when compared to 1D-RT models. Furthermore, the assumption is difficult to meet, particularly for structurally complex canopies like forest stands and sparse landscape. As mentioned before, the 1D-RTMs are based on the assumption that plant canopies are horizontally uniform turbid medium (Ross 1989). As as results, 1D approaches are usually used for mapping homogeneous grass and ignored for tree species discrimination.

This highlight the value of the development strategy that use 1D-RT models (Jacquemoud & Baret 1990) for modelling species distribution. Most IAPs grow in clusters, forming homogeneous landscape. Therefore, we argue that one-dimensional RTMs can be suitable in mapping distribution of IAPs. We, therefore, explored the utility of scaling *in situ* species leaf reflectance to canopy using 1-D RTM with the objective of discriminating invasive *Acacia mearnsii* from native species. Furthermore, we investigated the usefulness of canopy properties of the species retrieved with RTMs to map the species at the field scale. To accurately assess the potential of 1D model for species discrimination we also explored 3D RTmodel.

In a nutshell, this study aimed to:

(i) explore the utility of scaling-up *in situ* leaf reflectance to canopy reflectance for distinguishing *Acacia species* from native species;
(ii) assess the potential of radiative transfer model inversion methodology for automating mapping distribution of Acacia species based on Sentinel-2 based retrieved CCC and LAI parameters.
(iii) explore whether generated *LAI* and *CCC* maps can be usefully translated to the spatial distribution of *A. mearnsii, A.dealbata* and native species from the study site in Kwa-Zulu Natal, South Africa using multispectral Sentinel-2 image.

In this paper, we propose two successive approaches to assess the effectiveness of RTM for the discrimination of IAPs from native trees. In the first approach, we explore the ability of the extended leaf-level reflectance measurements of the species to the canopy scale for the discrimination of *Acacia* spp. from native trees. The study combined radiative transfer modelling with Analytical spectral device (ASD) leaf measurements to seek a way to overcome the problem of unavailability of the canopy–level hyperspectral data. We used canopy RT model to up-scale leaf optical properties to generate canopy reflectance for each IAP and native species. By upscaling measured leaf reflectance, the actual biochemical properties of the trees are used to constrain the model and makes the resulting canopy reflectance to mimic the actual canopy reflectance of the trees. The leaf reflectance was measured from experimental pot plants using a field spectrometer (400–2500 nm, 2100 bands), the ASD. The resulting simulated canopy spectra were then used to train and test a classifier, and the accuracy of this classifier was compared with that trained with measured canopy reflectance of the species.

The second approach presents the assessment of the classifier trained with simulations from 1D and the 3D-RT model and comparing the performance of the classifiers with that developed with the reflectance of the real Sentinel-2 image of the study area. Furthermore, we explored the LAI and CCC properties of the species retrieved using 1D and the 3D-RT model for mapping distribution of IAPs at a landscape scale and discussed the modelling results. Although simulations are accurate and can dramatically reduce field survey costs and increase time efficiency, the approach cannot replace experimental data for biodiversity assessment. The methodology of this paper is divided in two main sections. The first methods section describes the up scaling of leaf ASD based measured reflectance of the studied species to canopy reflectance using PROSAIL RT model and compared their usefulness with the canopy reflectance measured by an ASD for discriminating IAPs from native species.

2. Material and methods

2.1. Species discrimination based on up-scaling leaf-level reflectance measurements of the species to the canopy scale using canopy RTM

The leaf spectral were measured from the leaves collected from the canopies of the outdoor experiment with potted plants (Table 1) located on the campus of the Council for Scientific and Industrial Research (CSIR) for three months. The selection of the potted native plants was based on the field observations in the study area. The species mentioned above were found to be dominant and in the same area as *A. mearnsii.* The species show differences in their foliage structures as depicted in Table 2. Approximately 1-metre tall species were purchased from Nkosi Indigenous Plant Species Nursery based in KwaZulu-Natal, South Africa. The potted plants were left to grow outdoors on the campus of the CSIR for three months (1 September 2016 to 22 December 2016) before the start of the spectral data collection. The trees were randomly placed to allow the same distribution of energy and other resources. Due to the limited rainfall during this period, we watered the plants once a week from 1 September to 21 December 2016. The experiment took into consideration the soil type of study area to be surveyed for mapping species distribution at landscape level.

Table 1. Plant species used to explore spectral separability of *Acacia mearnsii* from native species.

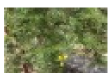
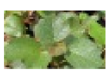
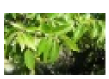
Species name	***Picture***	***Main characteristics***
Acacia mearnsii		Evergreen tree, 6–20 m high. Fast growing leguminous (nitrogen fixing) tree. The leaves are dark olive green branchlets with all parts finely hairy. Leaflets short (1.5–4 mm).Flowering (August-September)
Vachellia karroo		A deciduous tree (7–12 m).Leaves (pinnate leaflets).Flowering (November-April)
Vachellia xanthphloea		A deciduous tree (7–12 m).Leaves (pinnate leaflets).Flowering (November-April)
Euclea crispa		Variable short shrub to medium tree (8–20 m).Flowering (December-May)
Dombeya tilicea		Scrambling shrub tree (10 m). Spiralled, ovate leaves.Flowering March-August.
Dombeya rotundifolia		A small deciduous tree (5–10 m). Spiralled irregular lobed dark green leaves.Flowering (July- September)
Olea africana		Tree; grow up to 14 m tall in the forest.Opposite, decussate, shiny, leathery leaves.Flowering (October-January)
Celtis africana		A deciduous tree (30 m).Leaves (smooth and slightly leathery).Flowering (August -October)

Table 2. Parameterisation of the 4SAIL model used for upscaling *in situ* leaf reflectance and absorbance.

Canopy optical properties	Abbreviation	Units	Values
Leaf area index	LAI		0.1–3
Dry/Wet soil factor	psoil		measured spectral data
Hotspot parameter	H-spot		0.01
Solar zenith angle	tts	deg	0-90°
Observer zenith angle	tto	deg	15-75°
Relative azimuth angle	psi	deg	0
Distribution of leaf angles	LIDFa, LIDFb		30–90

2.1.1. Canopy level spectra measurements

We conducted nine sets of canopy radiance measurements on a biweekly basis from 23 December 2016 to 31 May 2017. They match the start and the end of the growing seasons in South Africa. The canopy measurements were carried out between 10:00 and 14:00 on cloud-free days. During the sampling campaign, we collected both leaf and canopy reflectance of the selected species between 350 and 2500 nm using an ASD FieldSpec FR 3 Spectroradiometer at one nm bandwidth (Analytical Spectral Devices Inc., Boulder, USA). We positioned the fibre optic (FOV 25°) at nadir and a height of approximately 30 cm above the undisturbed individual tree canopy. Fig. 1 summarises the methodology applied in this work.

Consequently, the field of view at the canopy level was circular, with a radius of 13.3 cm, and the field of view area was covered entirely by leaves to ensure standardised measurements. Furthermore, we eliminated the interference of the background (grass and bare ground) reflectance by placing the pot on a black sheet (board) (Cho *et al*. 2012). We used a ladder for the canopy measurements to ensure that the entire canopy was covered to account for the canopy spectral variability. In order to get the canopy spectral data of each plant, we randomly took six radiance readings and calculated the average to get one canopy reflectance. Overall we collected 80 × 6 measurements per survey. To reduce the effects of changing atmospheric and solar conditions, the reflectance of a Spectralon white reference panel was recorded every 10–15 measurements. The reflectance of the individual tree was converted using a reference measurement for each sample by dividing the reflected target radiance by the irradiance of the white Spectralon® panel.

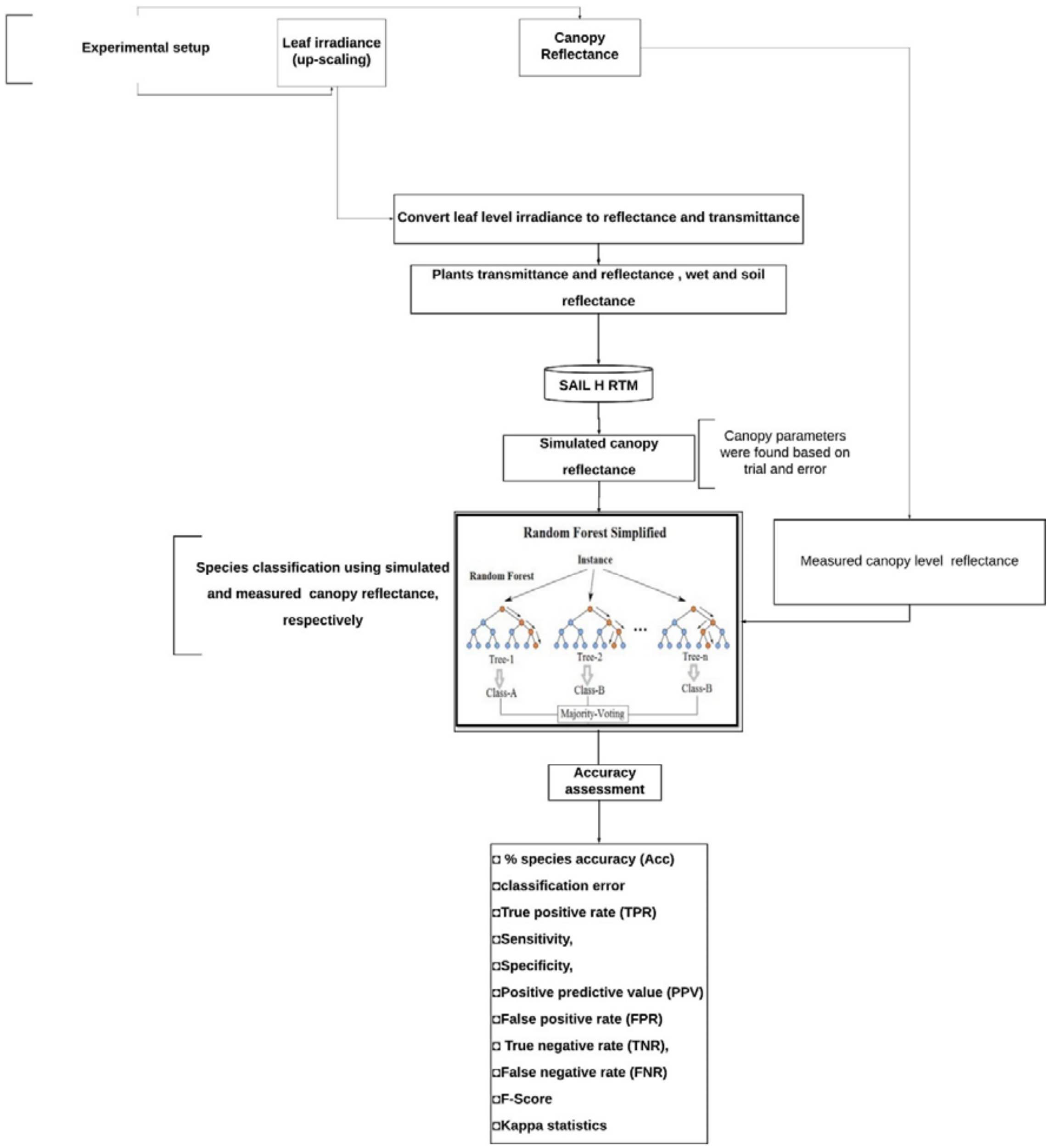


Fig. 1. Schematic illustration of the species discrimination approach using an RTM-based up-scaled leaf reflectance.

2.1.2. Leaf level spectra measurements

After every canopy measurement, ten leaves per tree were randomly harvested and taken to a dark laboratory room with the walls and the ceiling coated with the black material. The darkroom was used to ensure stable atmospheric and uniform illumination conditions (Clark et al. 2005; Ramoelo et at. 2011) We placed the leaves on a non-reflective black background surface to avoid the impact of external illumination. However, the leaflets of the pinnate leaves of *A. mearnsii, V. karroo* and *V. xanthophloea* (refer to Table. 1) were smaller than the sensor FOV, and their reflectance measurements would therefore not be truly representative of the leaf morphology and biochemistry because the leaf area per unit surface would also impact the measured reflectance. Consequently, we counteracted this effect by stacking the leaflets to simulate a continuous layer of leaves (Clark 2005). The fibre optic with a FOV of 25° was attached to the leaf clip and placed in a nadir position from approximately 4 cm above the leaves. However, in case of bigger leaves, no stacking of the leaves was done during spectra data collection. Capturing leaf spectral properties of the species, five measurements were collected per leaf by repositioning the leaf clip at five different positions for each scan. The reflectance of the individual plant was obtained by averaging the collected spectra reflectance per plant. The resultant spectral database included that of *A. mearnsii* and seven grown native tree species.

2.1.3. Spectral reflectance pre-processing

The pre-processing of spectral reflectance was conducted using the Field Spectroscopy Facility (FSF) Post-Processing Toolbox (MATLAB toolbox)**.**

(i) the toolbox allowed for the exclusion of outliers caused by measurement errors and atmospheric interference;
(ii) the 350–399 nm bands were not included in the analysis, thus limiting the spectral range to the traditional visible (VIS) to shortwave infrared (SWIR) (400 nm to 2500 nm);
(iii) in the case of canopy spectral reflectance, we removed SWIR ranges with high noise that were identified through literature and visual inspection, that is, 1350–1460 nm and 1790–1960 nm (Große-Stoltenberg, 2016).

After applying the pre-processing, 1759 canopy bands were left for analysis. Lastly, we eliminated sensor noise by using a moving Savitzky-Golay filter (Savitzky and Golay, 1964) with nine-point window size and second polynomial order. Furthermore, we explored various spectral transformation algorithms to evaluate the impact of spectral transformation on species discrimination. We considered the following methods: multiplicative scatter correction (MSC) standard normal variation (SNV) and first derivatives. The performance of the transformed data was then compared with that of the untransformed spectral dataset.

At each site, upper-canopy leaf samples were collected using shotgun, pole-clipping and tree-climbing techniques as necessary. Species were carefully selected to control for full sunlight canopies. This process requires that two or more trained workers agree that at least 50% of a selected canopy maintains an unobstructed exposure to the sky. Individual canopies meeting this criterion were then marked, and a voucher specimen was collected.

2.2. Radiative transfer modelling, up-scaling and forward simulation

Canopy scale radiative transfer models (4SAIL2) was used in this study, describing species at the canopy scale using measured leaf optical properties. Simulation of canopy characteristic of each species used the novelty of integrating measured leaf reflectance and variables mandatory for parameterising 4SAIL2 provided in Table 2. The scaling of the 4SAIL2 was performed by ranging variables values, which is a way to include variation in canopy structure and geometry of species canopies (Table 2). By adjusting parameter ranges on simulations, actual species reflectance can be successfully mimicked from simulated data. The mean and standard deviation leaf reflectance of each species was used together with canopy parameters to randomly simulate canopy reflectance. We first performed trial-and-error experiments to test the effect of suitable parameter ranges for simulating canopy reflectance from the leaf reflectance data. After the test, we found that an LAI of between 0.1 and 3 was adequate for simulating canopy spectral reflectance that matches that of the native species. Therefore, LAI values were ranged from one to three; we parameterised the leaf angle in such a way that accommodates all species leaf types. The hotspot effect in plants was fixed based on prior knowledge. The diffuse radiation was between 10 and 50 (Table 2), and both dry and wet soil reflectance was collected during spectral data collection campaigns. Finally, we used solar zenith and observer zenith angles ranging from 0 to 90° and 15–75° respectively. The rest of the parameters were fixed (Table 2). The model was allowed to run in a forward mode and then in the inverse model using measured canopy reflectance data. The resulting canopy data was used to train a classifier for discriminating *Acacia mearnsii* from native species. The accuracy of the classifier training and tested with simulated data was then compared with that of classifier trained using measured canopy data of the species.

2.3. Developing a classifier to perform discrimination of A. Mearnsii from native trees

The tree-based random forest discriminant analysis (RF-DA) methods (Jones 2015; Lemmond 2008) has been used for species. RF-DA was used because it has been shown adequately for species discrimination using remote sensing data, particularly forests landscape. Fig. 1 summarises the methodology applied in this work. The RF-DA used was adapted from a Matlab Fathom Toolbox code provided by Jones (2015). For comparable results, the measured and simulated datasets had the same number of samples. Both simulated and measured data based models used 70% of the simulated canopy reflectance for training and 30% for validation. The prediction accuracy of the species was assessed through Multi-Class Confusion Matrix. The matrix provides the prediction statistics parameters of each class and for the overall classification. The statistical parameters used to assess accuracy were species-specific accuracy, separability error, sensitivity, specificity, precision and FalsePositiveRate, F1-score and kappa coefficient. These metrics provided the detailed performance of the classifier and were recommended for species discrimination model assessment (Fielding and Bell 1997; Lurz *et al*. 2001).

3. Application of a radiative transfer model to discriminate IAPs distribution from sentinel-2 image at field scale

Here we develop a RT model inversion methodology for physically based mapping of *A. mearnsii* based on spectral data, LAI and CCC retrieved using Sentinel-2 data (10 bands with 20 m of spatial resolution). Most importantly, we explored the benefit of retrieving tree LAI and CCC parameters without priori field data. The study explored coupled leaf (PROSPECT-5) and canopy (4SAIL2) models to map distribution of invasive alien *Acacia* species.

3.1. Study site

We collected species geographical location of the species in sub-montane forests of uThukela District Municipality in KwaZulu-Natal Province, South Africa (Fig. 3). To be specific, the area is near Van Reenen's Pass (Lat-28.488023°and Lon 29.301116°) on the Great Escarpment of the Drakensberg. The closest villages are Geluksberg, Howe, Nqula and Wittekop. The closest towns are Bergville and Harrismith.

The primary uses of land are ranching and agriculture. The area is dominated by native tree species *Celtis africana, Dais cotinifolia, Diospyros lycioides, Podocarpus latifolius, Searsia Rehmannia, Senegalia caffra, Vachellia sieberiana, Peltophorum africanum* and *Leucosidea sericea.* The invasive species of interest was *A. mearnsii*. However, during the field survey, *Acacia dealbata* (silver wattle) was found to invade the study area. *A. dealbata* is the second most aggressive invasive species after *A. mearnsii* in South Africa. Besides, *A. mearnsii* is listed among the 100 most aggressive invaders in the world (Global invasive species database, "One Hundred of the World's Worst Invasive Alien Species.

Morphologically, *A. mearnsii* and *A. dealbata* are similar. They are both fast-growing evergreen leguminous trees that can grow up to 30 m high. Their leaves are bipinnate, but with different colour and texture. For example, *A. mearnsii* has finely hairy dark olive-green leaves, whereas the leaves of *A. dealbata* are blue-green to a silvery grey and broad. The greyish silver leaves of *A. dealbata* are the main factor that distinguishes it from *A. mearnsii*. Other than the leaves, the species are more easily distinguished during the flowering season. The flowers of *A. mearnsii* are pale yellow and spherical, whereas *A. dealbata* has bright yellow with globe-shaped heads. In the study area, *A. dealbata* start to bloom from June to August, while the flowers of *A. mearnsii* start to appear in August and peak in September and October (Impson et al., 2008).

3.2. Species sampling

The sampling method adopted in the study was transecting survey. This method was adopted to detect the invasive plants along elevation gradients. Some transects started along the rivers and ended on top of the hill far from the wet area. Transects were created by overlaying Sentinel-2 imagery over the study area on Google Earth. Twenty transects were created based on the Sentinel-2 MSI 20 m pixel size. The length of the transects was determined based on density and homogeneity of the tree canopies and varied between 50 m and 200 m in length. Longer transects were created for areas with a sparse and high variation of vegetation cover, while shorter ones were for dense and homogeneous cover. Coordinates at the start and end of the transects were uploaded into a handheld Garmin GPS device, which was used to navigate to the area in the field. In the field, tree species along the transects were identified with the help of local species expert and coordinates were recorded using a GPS device.

During data collection, canopies of species that overlapped one another were sampled separately. Standing dead canopies were not counted but were noted to avoid confusion during classification. Dense patches of *A. mearnsii* and *A. dealbata* scattered throughout the study area were recorded. A summary of the sampled trees is presented in Table 3. From the desktop, GPS points collected from the field were converted to Keyhole Markup Language (kml) file format and overlaid on Google Earth. This was done to validate that the GPS points correspond to each sampled tree. Furthermore, we manually polygonised *A. mearnsii*, *A. dealbata* and un-infested (natural forest) clusters in Google Earth.

Table 3. The number and name of the sampled native and invasive alien trees used for training and validating classification model.

Dominant tree species	Number of sampled tree per species
Accacia gerrardii	180
Acacia dealbata/silver wattle	150
Acacia mearnsii/black wattle	166
Podocarpus latifolius	89
Vachellia sieberiana	101
Diospyros lycioides	33
Senegalia caffra	50
Leucosidea sericea	30
Peltophorum africanum	40
Celtis africana	69
Searsia rehmannia	75

3.3. Acquisition and pre-processing of the Sentinel-2 imagery

The Sentinel-2 Multispectral Instrument (MSI) images (tile number-L1C_T35JQJ) with zero to less than 10% cloud cover was downloaded from the United States Geological Survey through the Earth Explorer search interface. The georeferenced images were March 2018, therefore, covering peal production season in South Africa. We processed reflectance images from Top-Of-Atmosphere (TOA) level to tree canopy reflectance level using iCOR (VITO 2017), available as a plug-in on the Sentinel Application Platform (SNAP) v5.0. According to VITO (2017) iCOR, correct Sentinel-2 MSI data by identifying water and land pixels using MODTRAN 5 radiative transfer model Look Up Tables (Berk et al., 2006). More information on the procedure is outlined in VITO (2017). The spectral band's characteristics of Sentinel-2 MSI images are presented in Table 4. The study site is mountainous; as a result, Sentinel Topographic Illumination Correction was performed. The correction was executed based on pixel-based Minnaert Correction Method (Ge et al., 2008), solar angles from scenes metadata and 10 m Shuttle Radar Topography Mission (SRTM) Digital Elevation Model downloaded from Google Earth Engine operating system.

Table 4. Sentinel-2 Multispectral Instrument (MSI) spectral characteristics: band centre and spatial resolution of the ten bands used for the inversion and discrimination of *Acacia mearnsii* from native trees.

Band name	Band width(nm)	Band centre(nm)	Spatial resolution(m)
Blue	96	490	10
Green	45	560	10
Red	39	665	10
Red-edge1	20	704	20
Red-edge2	18	740	20
Red-edge3	28	783	20
Near-Infrared-1	141	842	10
Near-Infrared-2	22	865	20
ShortwaveInfrared-1	142	1610	20
ShortwaveInfrared-2	240	2190	20

3.4. PROSAIL model inversion to derive LAI and CCC

3.4.1. Ill-posed problem of RTM

Retrieval of vegetation parameters from remote sensing observations using inversion of RTMs is an ill-posed problem. The solution tends to be the same for all situations (Combal et al., 2002). Regularization of ill-posed problems with ancillary information, spatial data as well as temporal constraints was shown to be effective (Dorrigo et al., 2008; Lauvernet et al., 2008). Because the models have not been implemented for distinguishing IAPs from native species on a landscape, no ancillary information can be confidently used in this study. This study will adopt retrieval approach proposed by Verrelst et al. (2012). The approach used classified species maps to retrieve the parameter of the species. The approach has been providing a good overall estimate of the parameters (Verrelst et al., 2012). However, the approach is useful when species or land cover is known. Briefly, the approach is another way of constraining RT models, subsequently solving the ill-posed problem. In this study, we developed a species distribution map before forward and inversion approaches of one-dimensional PRO5SAIL4 canopy radiative transfer model. Therefore, the research presented in this paper took the RTM inversion approach presented in Verrelst et al. (2012) and adopt it for the discrimination of IAPs (*Acacia spp*.) from native trees using Sentinel-2 images as the sole source of information. Fig. 2 outlined the methodology developed for this study. The approach is based on LUT-species distribution image-based inversion strategies demonstrated in Verrelst et al. (2012). The strategy is useful in the absence of *in situ* parameter estimates, such as in this study.

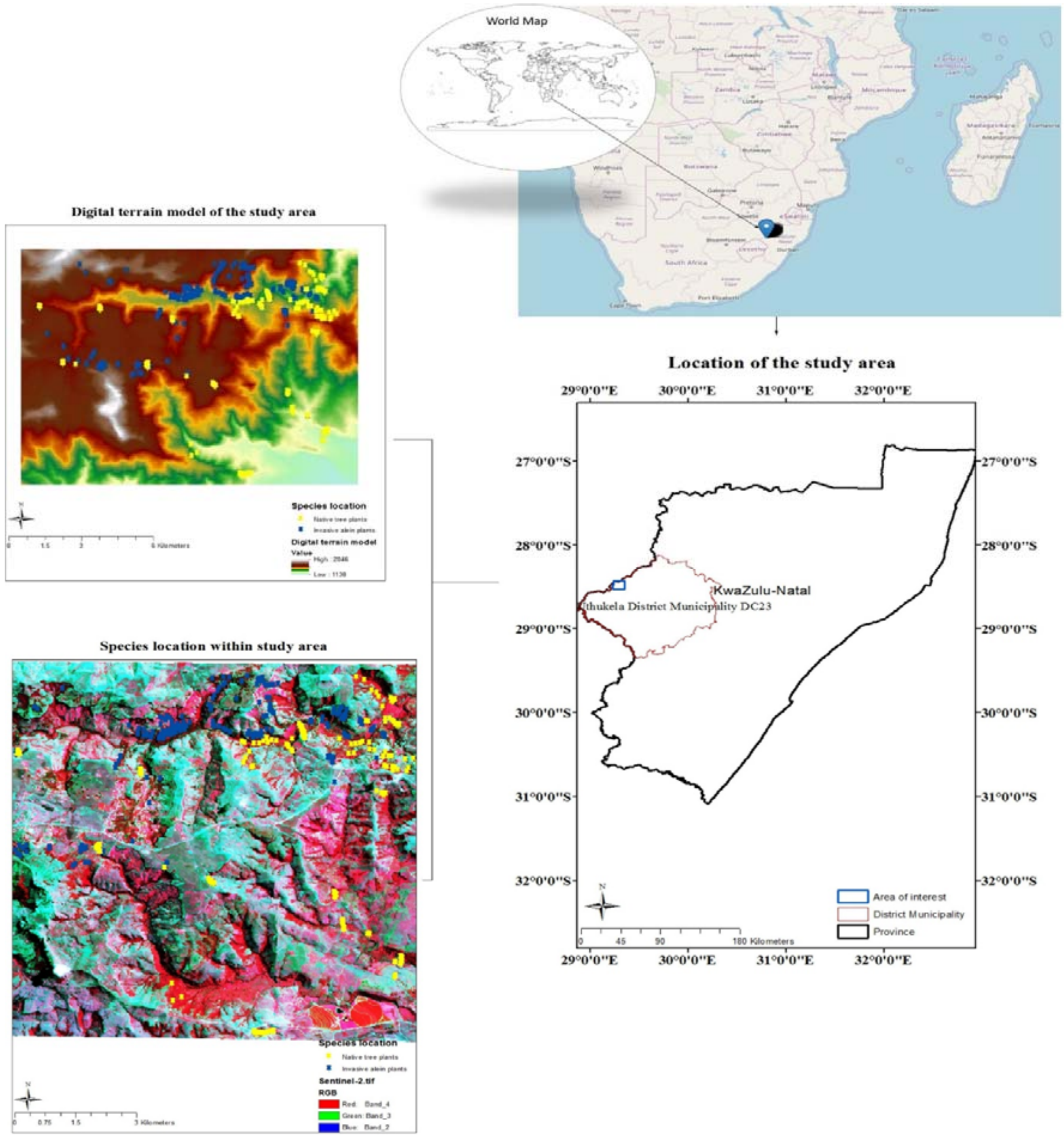


Fig. 2. The study is relative to work and South Africa.

3.4.2. Classification

First, we distinguished invasive species from native species using field species data and the discriminant random forest (DRF) methodology (Lemmond et al., 2008). The DRF is an ensemble of decision trees that leverages linear discriminant analysis to perform node splitting and determination of optimal linear decision boundary (Lemmond *et al*. 2008). Like conventional random forest (RF), DRF uses bagging and random feature selection approaches to select essential variables for the classification problem. The model measures variable importance and feature ranking using a Mean Decrease in accuracy, through random permutation. The most important aspect about RF is its insensitivity to the number of input data used. As a result, it has been widely used for vegetation classification using satellite data (Naidoo et al., 2012; Lehmann et al., 2016). The robustness could be attributed to the fact that RF is not affected by distribution (Fu et al., 2012) and redundancy of the data. However, in cases of unbalanced class samples, RF, like most classifiers, often exhibits a predictive bias in favour of oversampled class (es). In this study, we circumvented implementing random oversampling examples (ROSE) (Lunardon et al., 2014) to the minority class. The advantage of the technique is that no information is lost during processing while improving prediction accuracy. The function resamples the minority class until it reaches the number of samples of the majority class. In this study, the number of trees to be grown was determined using a grid search approach. In this data, we implemented the model in Python with Scikit-Learn (Pedregosa et al., 2011). During model building, a subset from the calibration dataset is selected using the bootstrap sampling approach and a classification tree is built. Each node within the tree is constructed by selecting a random subset of the predictors (i.e., Sentinel bands) and determining which predictors provide the most effective split for maximising purity in the resultant groups. Nodes are continuously added to the tree until the desired number of trees have been built (*ntree*). The algorithm performs the prediction accuracy of each forest tree and predictors with the most votes are selected for tree species prediction.

3.4.3. FLIGHT and PROSAIL model inversion to derive LAI and CCC of the trees

We simulated canopy reflectance corresponding to Sentinel-2 band centres. To investigate the potential of mapping *A. mearnsii* at the landscape level using Sentinel-2 MSI data the respective pre-defined spectral response functions provided incorporated in ARTMO was used to convert hyperspectral bands into Sentinel-2 MSI bands centres. The resultant centre wavelengths were: Sentinel-2 (490 nm; 560 nm; 665 nm; 705 nm; 740 nm; 785 nm; 842 nm, 1601 nm and 2190 nm). The simulations were performed using the combined PROSPECT-5 leaf optical properties model (Féret et al. 2008) and SAIL canopy bidirectional reflectance model (Verhoef *et al.* 2007). The PROSPECT-5 is an extension of the PROSPECT leaf optical properties spectra model by Jacquemoud and Baret (1990). PROSPECT-5 simulates reflectance and transmittance from 400 to 2500 nm spectrum region at a 1-nm spectral sampling interval. The model simulates the plant leaf reflectance and transmittance using leaf optical properties such as leaf mesophyll structure index (N), leaf chlorophyll content (C_{ab}), leaf dry matter content (C_{dm}), leaf water content (C_w) and leaf brown pigment content (C_{bp}) (Eq. (1)). On the other hand, 4SAIL canopy bidirectional reflectance simulates top-of-canopy reflectance using PROSPECT outputs (i.e. leaf reflectance and transmittance) as well as LAI, the average leaf angle (ALA), hotspot parameter, a wet and dry soil reflectance, geometrical illumination and view information of the remote sensing image to be used for inversion of the model (Eq. (2)).

$$LRT = f(N, Cab, Car, Cbrown, Cw, Cm) \tag{1}$$

$$PROSAIL_{rdot,rsot,rddt,rdt} = f(PROSPECT, LIDFa, LIDFb, LAI, hspot, tts, tto, psi, rsoil) \tag{2}$$

Table 5. Range and distribution of input parameters of the PROSPECT-5 and 4SAIL radiative transfer model used to generate the training database corresponding to Sentinel-2 MSI bands configurations.

Parameter	Abbreviation	Unit	Ranges
Leaf parameters-PROSPECT-5			
Leaf chlorophyll *a* + b concentration	Leaf Chl (a and b)	$\mu g/cm^2$	1–300
Leaf dry matter content	Cm	g/cm^2	0.01–0.05
Leaf water thickness	Cw	cm	0–0.005
Leaf mesophyll structural parameter	N	unitless	1–3
Carotenoid	Car	$\mu g/cm^2$	10–50
Brown pigment	C_{brown}	Arbitrary	10–50
Canopy structural variables-4SAIL			
Leaf area index	LAI	(m^2/m^2)	1–6
Dry/wet soil	soil		Soil reflectance from Sentinel-2
Hot spot	Hspot		0.01
VIEW AND ILLUMINATION GEOMETRY			
Sun zenith angle,	tts	deg	$15\text{-}90^0$
Azimuth angle	psi	deg	0
Sensor viewing angle	tto	deg	15–75

3.4.4. Parameters range and distributions used to create the synthetic top-of-canopy reflectance database

The minimum and maximum, as well as distributions of the other PRO5SAIL4 parameters, are shown in Table 5. Parameters constraints were taken from other studies working in the sites in which evergreen and deciduous trees are mixed. Furthermore, the trial-and-error approach was used to find parameter bounds that produce simulated spectra with the smallest RMSE to the measured spectral reflectance. It should be noted that, in this study, we parameterised species groups separately; that is, for each species class (IAPs and native plants species). Distribution of the parameter varied, for instance, for LAI and leaf chlorophyll (Cab), the content Gaussian distribution method was adopted to emphasise variability in the biochemical and biophysical properties of the species. Hotspot parameter was parameterised according to studies conducted on forested areas (e.g.Jacquemoud et al., 2009). We extracted geometric and viewing information from Sentinel-MSI metadata. The soil reflectance was also extracted from bare dry and wet areas respectively within the study area (see Fig. 3).

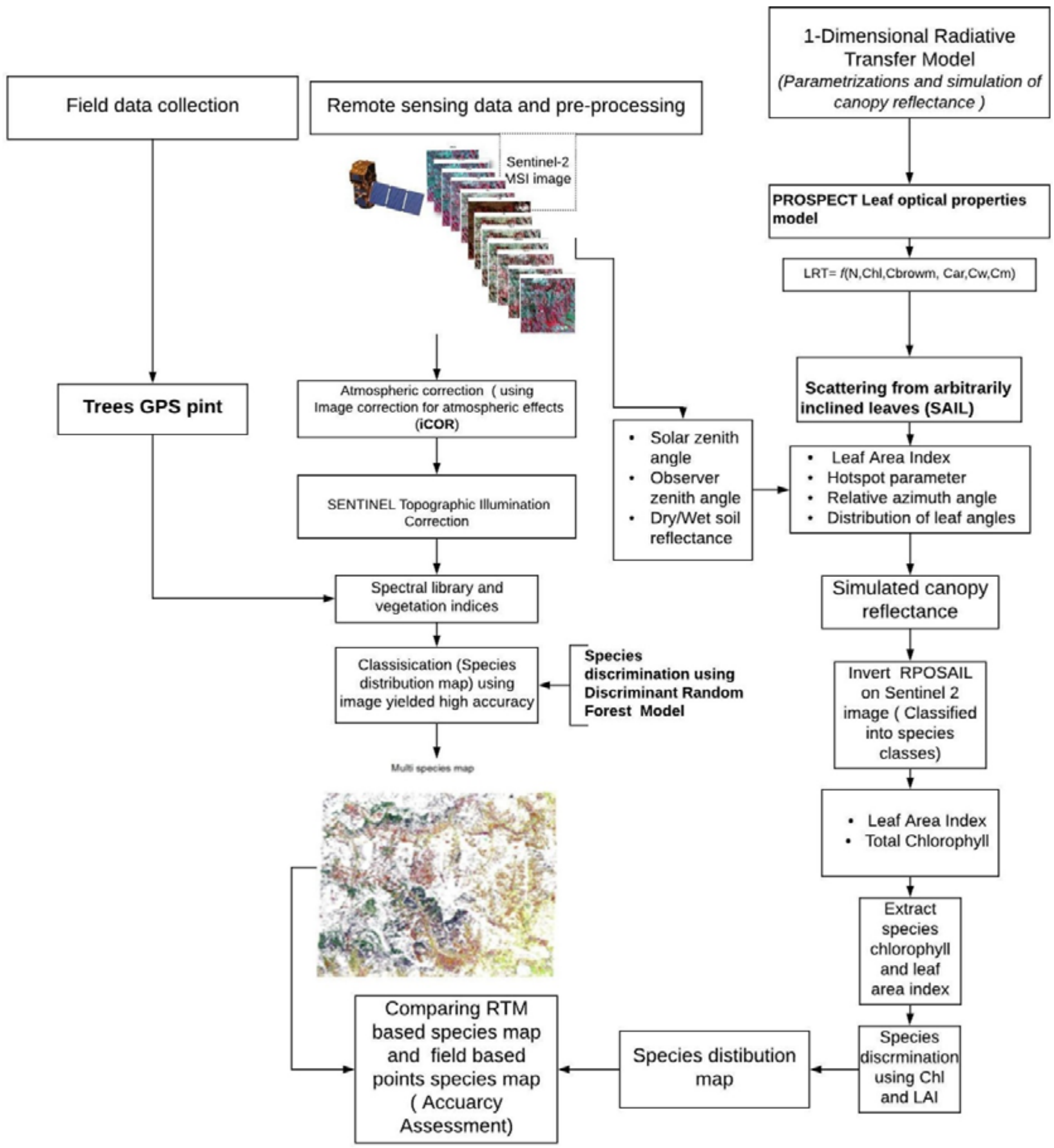


Fig. 3. Schematic illustrations of the species discrimination approach using a radiative transfer model.

3.4.5. Simulation of canopy reflectance

The parameter value ranges presented in Table 5 used to simulate Sentinel-2 canopy reflectance in ten bands. According to Combal et al. (2002), the size of the simulated reflectance database depends on the complexity of the problem. The study suggested that a more complex problem requires a large look-up table of simulated reflectance. The less complicated problem can be solved with close to 10 000 cases (Combal et al., 2002). In this study, we have randomly simulated various sizes of LUT (i.e. 10,000, 30,000, 50,000, 70,000, 90,000 and 120,000), but LUT of 50,000 top-of-canopy reflectance cases were found to be sufficient for this study species distribution modelling problem.

3.4.6. PROSAIL Sentinel-2 MSI-based inversion model for the species CCC and LAI retrieval

In this study, the PROSAIL model inversion was applied according to the method proposed in Verrelst et al. (2012). The inversion of this model against measured optical information of vegetation allows retrievals of both CCC and LAI at the sensor pixel scale. The inversion approach adopted here used class-based inversion approach (Verrelst et al. 2012), in which species distribution maps developed in 3.5.2 was used as a base map for inverting both 1D-PROSAIL and 3D Flight models (Verrelst et al. 2012). The base map consisted of two classes, i.e., native and Acacia tree species, which was used to parameterize and invert the models. By doing so, each species canopy was converted into LAI and CCC, respectively. The outputs were LAI, and CCC distribution maps of the species (IAPs and native species) for the study area showed in Fig. 1. We executed the retrieval of the species parameters into an Automated Radiative Transfer Models Operator (ARTMO) toolbox (Verrelst *et al*. 2011). The toolbox is the Matlab-based graphical user interface (GUI). Subsequently, LAI and CCC corresponding to the species geographical location were extracted and used to discriminate IAPs from native plants species using DRF algorithm.

4. Results

4.1. Upscaling leaf-level reflectance to the canopy reflectance using the 4SAIL canopy radiative transfer model

4.1.1. Parameters sensitivity analysis

Canopy reflectance changes with changing LAI (Fig. 4) and illumination geometry (*tts*) (Fig. 4). Fig. 4 and Fig. 5 showed that canopy reflectance of the species could be successfully modelled with parameter LAI and illumination geometry ranging between 1 and 3 and 55 to 85°respectively. However, LAI value showed to produced simulated reflectance similar to that measured with N ranging between 1 and 2 and planophile leaf type values. Fig. 4 shows the mean reflectance for the measured and simulated species spectral data.Fig. 6.

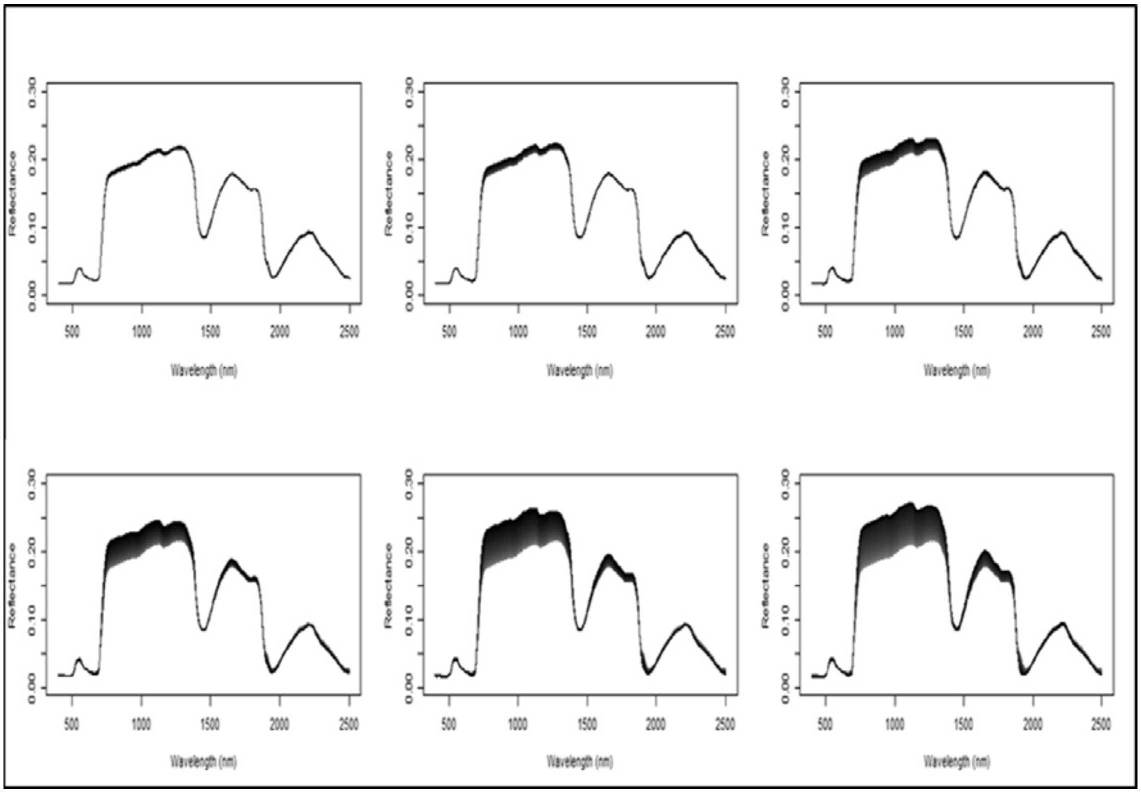


Fig. 4. Depiction of the effect of the solar zenith angle simply plotting the spectra in different.

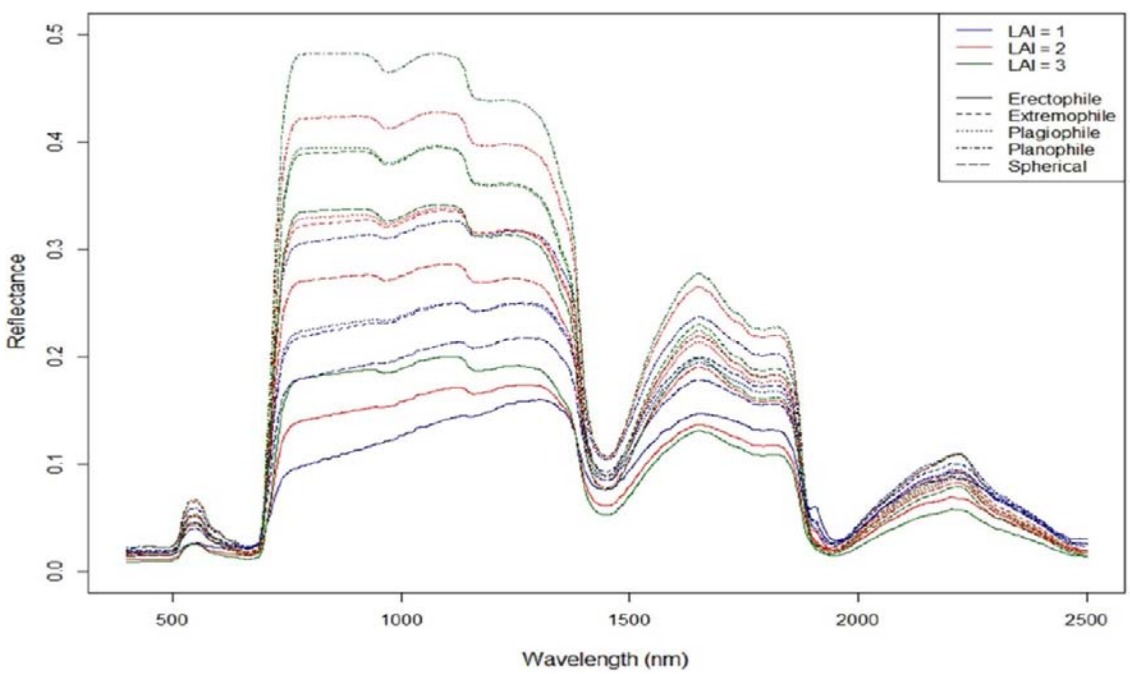


Fig. 5. Presenting the effect of LAI and leaf angle distribution parameters on the reflectance values of the species.

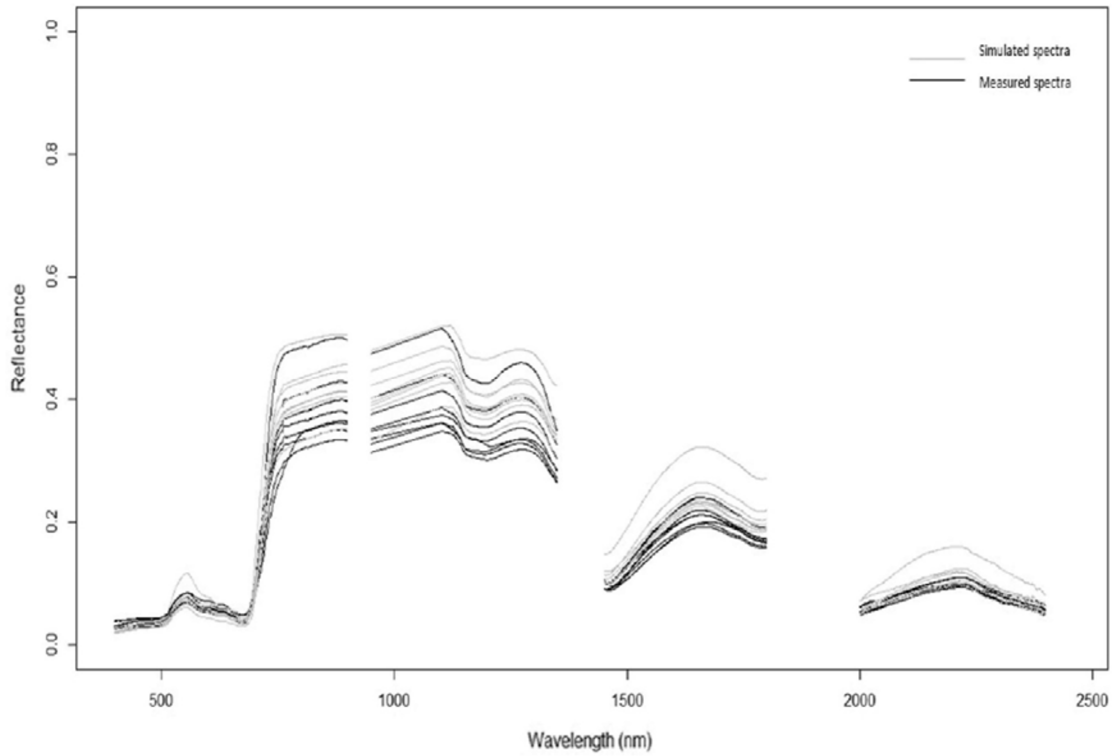


Fig. 6. Comparison between the mean of simulated spectra and measured canopy spectra of *A. mearnsii* and seven native species using LAI values ranging from 1 to 3, N ranged from 1 to 2 and illumination geometry ranged from 55 to 85°. Grey = simulated spectral reflectance; Black = measured spectral reflectance.

4.1.2. Species discrimination using simulated canopy reflectance

The experimental and simulated reflectance spectral yielded similar discrimination accuracies for *A. mearnsii*. The accuracies for *A. mearnsii* ranged from 79% to 0.92 and 80%, and 100% for the experimental and simulated data, irrespectively of classifier (iECV-DA and DRF) (Table 6).

Table 6. Classification accuracies of *Acacia mearnsii* using measured and RTM simulated data.

Statistic accuracy metrics	iECVA-DA	RF-DA	iECVA-DA	RF-DA	iECVA-DA	RF-DA	iECVA-DA	RF-DA
	ASD fieldspec measured canopy spectra				PROSAIL RTM simulated canopy spectra			
	Acacia mearnsii		Overall accaracy		*Acacia mearnsii*		Overall accaracy	
Visible region (400–650 nm)								
Accuracy (%)	85.7	100	66.25	60.7	85.7	91.8	66.7	54.6
Specificity (%)	94.5	100	95.2	94.3	94.4	98.5	67.5	94.2
Prediction power (%)	85.7	100	66.25	60.7	85.7	91.8	66.7	54.6
Kappa	0.69	0.96	0.62	0.6	0.69	0.84	0.62	0.51
Red-edge region (651–750 nm)								
Accuracy (%)	90	100	63.75	61.5	91	91	63.75	66.67
Specificity (%)	95.7	100	94.86	97.2	100	100	64.89	94.4
Prediction power (%)	90	100	63.75	61.5	91	91	63.75	66.67
Kappa	0.89	0.95	0.59	0.7	0.86	0.85	0.59	0.58
Near-infrared region(751–1300 nm)								
Accuracy (%)	90	100	56.25	66.7	92.78	100	56.25	57.14
Specificity (%)	98.5	100	93.8	94.4	95.8	100	55.61	96.97
Prediction power (%)	90	100	56.25	60	92.78	100	56.25	57.14
Kappa	0.84	1	0.5	0.58	0.7	1	0.5	0.62
Early-shortwave region (1301–1460 nm)								
Accuracy (%)	89.7	100	58.7	55.6	87.7	100	58.7	55.6
Specificity (%)	97.1	98.6	94.2	92.9	95.8	98.9	58.2	92.9

Prediction power (%)	89.7	100	58.7	55.6	87.7	100	58.7	55.6
Kappa	0.82	0.94	0.52	0.46	0.7	0.94	0.52	0.46
Mid-shortwave region (1461–1900 nm)								
Accuracy (%)	88.89	100	68.75	71.43	80	90	68.7	71.4
Specificity (%)	97.18	98.59	95.54	93.6	97.2	98.59	71.4	93.2
Prediction power (%)	88.89	100	68.75	71.43	80	90	68.7	71.4
Kappa	0.82	0.94	0.65	0.55	0.82	0.94	0.65	0.55
Far-shortwave region (1901–2500 nm)								
Accuracy (%)	77.7	75	50	66.67	85	87.2	50	63.6
Specificity (%)	95.8	98.53	92.9	91.89	94.44	98.5	49.7	95.65
Prediction power (%)	77.9	75	50	40	85	87.2	50	63.64
Kappa	0.76	0.79	0.43	0.47	0.63	0.79	0.43	0.62
Global model (400–2500 nm)								
Accuracy (%)	95.43	100	57.5	61.54	96	100	57.5	61.54
Specificity (%)	93.15	100	94.02	97.02	94.2	100	55	97
Prediction power (%)	95.43	100	57.5	80	80	100	57.5	61.5
Kappa	0.55	1	0.5	0.65	0.65	1	0.5	0.65

4.2. Sentinel-2 MSI image-based species discrimination

4.2.1. Simulation of the canopy reflectance based on Sentinel-2 MSI bands

The following leaf input parameters ranges were used to simulate Sentinel-2 canopy reflectance using PROSAIL: N = 1.4 (fixed), Cab = 60–100 ug/cm^2 (ranges increment of 5), Car = 10, C_{brown} = 0, C_w = 0.01, and Cm = 0.01. The same canopy parameters used in the previous section were used for the simulation of Sentinel-2 canopy reflectance (Fig. 7).

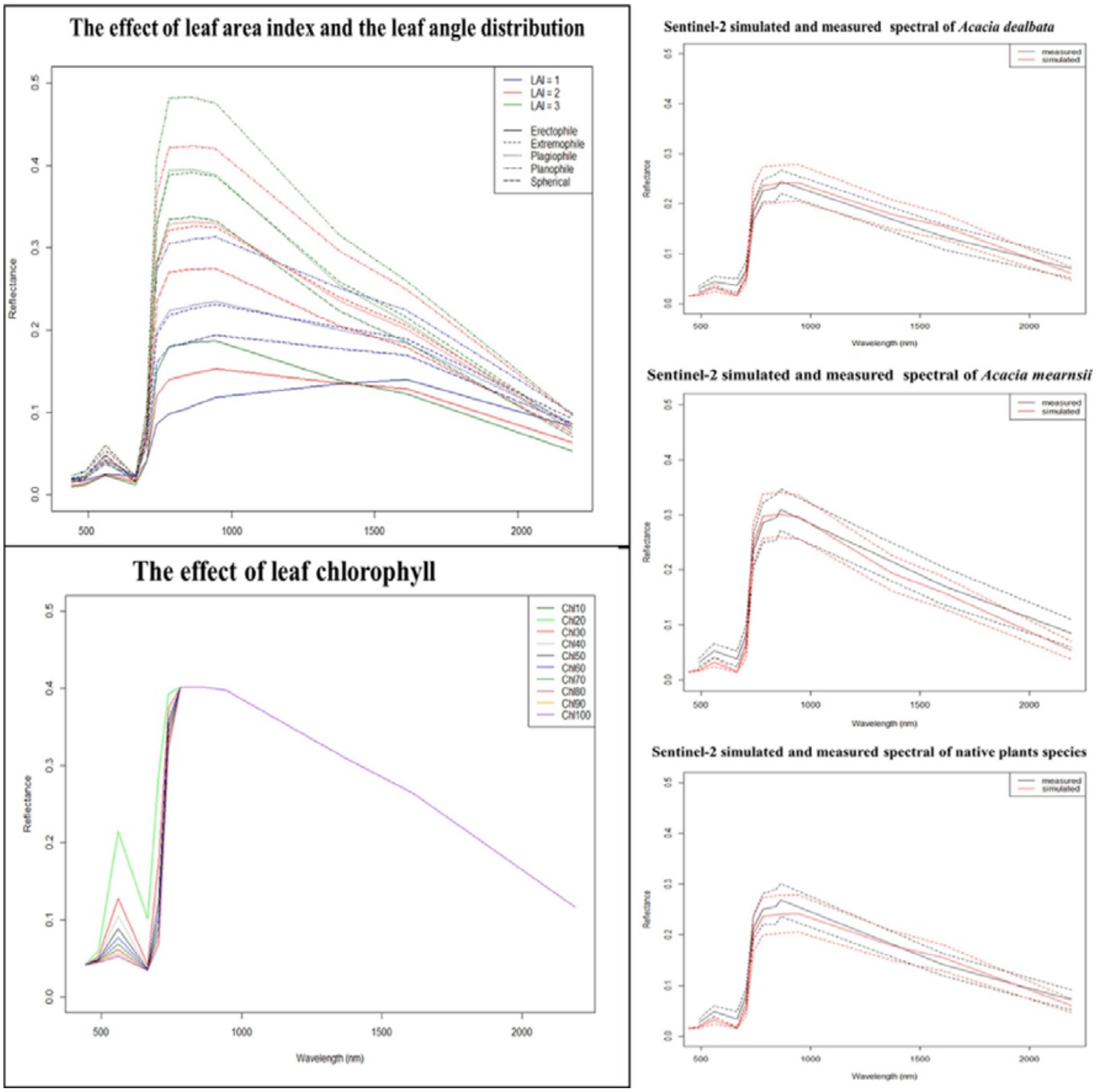


Fig. 7. The effect of varying leaf and canopy parameter of the simulated reflectance spectral data.

4.2.2. PROSAIL Sentinel-2 image discriminatory power for IAPs from native species from the study area

Our results showed that discrimination of IAPs from native species using simulated Sentinel-2 MSI reflectance was slightly lower (Accuracy = 82.2% and kappa coefficient = 0.80) compare to the field based sampling classification approach (Accuracy = 90.2% (IAP), 82.2% (NAT) and kappa coefficient = 0.84 (IAP), 0.78 (NAT)) (Table 7). On the hand, *A. mearnsii* was successfully discriminated from native species using LAI and CCC maps derived created from the study area using PROSAIL inversion on Sentinel-2 image (Fig. 8). These maps showed considerable high values of LAI and CCC for the IAPs (*A. mearnsii* and *A. dealbata*) compared to native species. Table 7 also report on the discrimination accuracies of IAPs (*A. mearnsii* and *A. dealbata*) from native species based on retrieved CCC and LAI of the species. *A. mearnsii* was successfully discriminated from native species and *A. dealbata* with accuracies

(% accuracy = 92.8%, 91.4% for CCC and LAI, respectively) (Table 7). According to our results, *A. dealbata* was misclassified as native species for both LAI and

Table 7. Confusion matrix and model assessment metrics for the species classified with discriminant random forest and different datasets.

Species	DRF-Cross -validation confusion matrix			Statistic metrics to describe model performance				
Discrimination using Sentinel-2 reflectance								
	AD	**Nat**	**AM**	**Accuracy (%)**	**Sensitivity (%)**	**Specificity (%)**	**Precision (%)**	**Kappa (%)**
AD	82.2	9.4	8.4	82.2	77.02	91	82.2	0.78
Nat	17.8	73.4	8.8	73.4	86.36	88	73.4	0.68
AM	6.4	3.4	90.2	90.2	84.56	95	90.2	0.84
Overall accuracy				80.2	94.56	96	80.2	0.78
Discrimination using forward PROSAIL RTM based simulated reflectance								
AD	77	14	9	77	90.02	91	77	0.72
Nat	19.6	70.2	10.2	70.2	86.36	88	70.2	0.75
AM	9	9	82.2	82.2	95.56	95	82.2	0.8
Overall accuracy				78.2	90.26	95	78.2	0.75
Discrimination using LAI retrieved using PROSAIL RTM								
AD	79.8	10.8	9.4	79.8	90	88	79.8	0.76
Nat	20.2	73.2	6.6	72.6	87	73	72.6	0.68
AM	5.4	2.8	91.4	91.4	96	91	91.4	0.82
Overall accuracy				89.3	94	95	89.3	0.8
Discrimination using LAI retrieved using PROSAIL RTM								
AD	82.8	9.4	7.8	82.8	83	86	91	0.84
Nat	5.8	90.6	3.6	90.6	91	90	95	0.75
AM	6.4	0.8	92.8	92.8	93	88	96	0.85
Overall accuracy				88.2	90	90	88.2	0.79

CCC based classification. The accuracy assessment showed acceptable overall accuracy ranging between 75.6% and 82.36% and a kappa statistic ranging between 0.68 and 0.78 (Table 7). Fig. 9 demonstrated distribution patterns of IAPs and native species based on field sampling data and class-based inversion of 1D- PRO5SAIL4 RTM as well as LAI and CCC species parameters. Overall, PRO5SAIL4 RTM based discrimination of *A. mearnsii* from native species confirmed acceptable overall accuracy using all datasets and showed distribution patterns of the species consistent with that observed in the field.

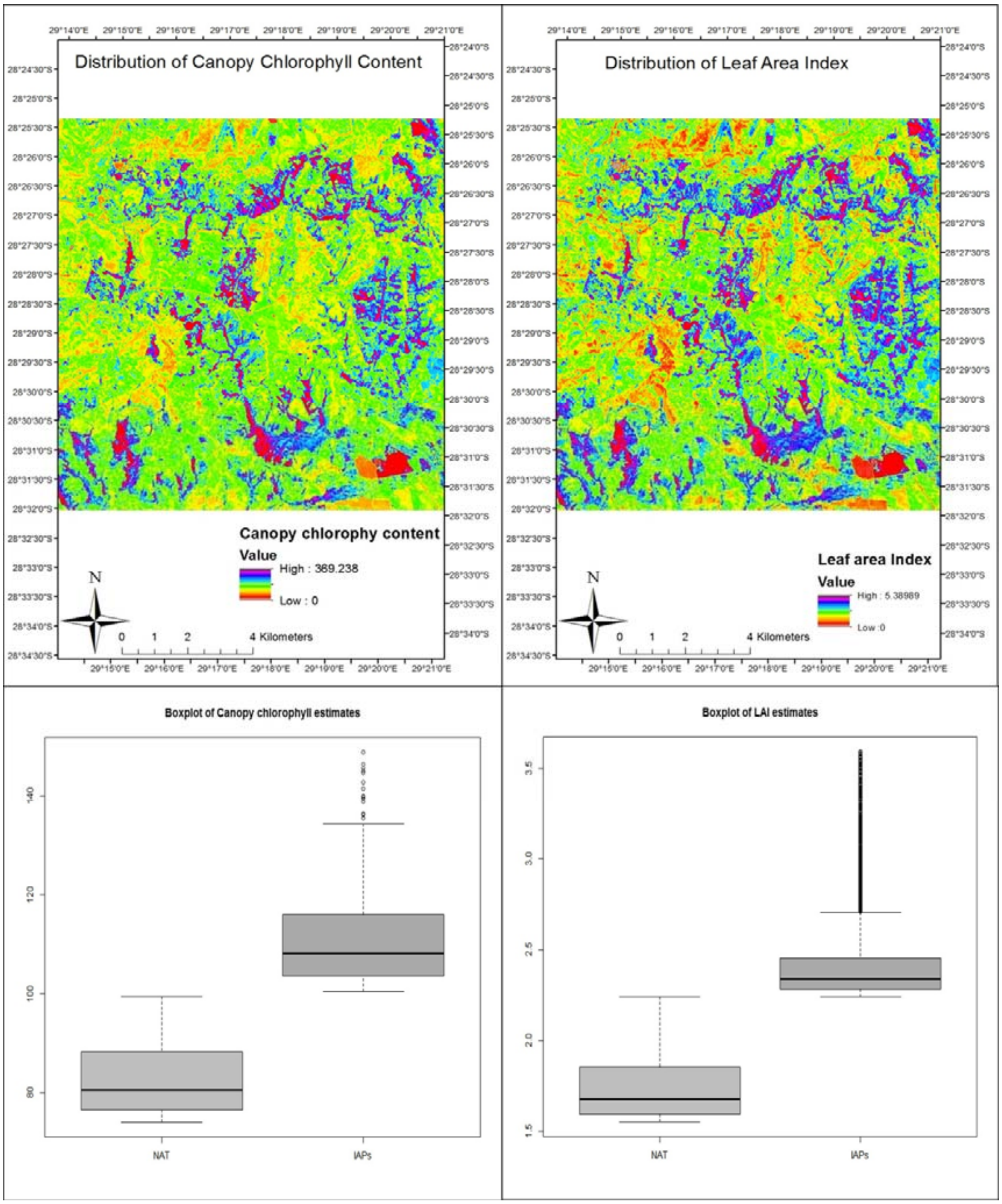


Fig. 8. Maps of alien invasive species (*A. mearnsii and A. dealbata*) produced using LAI and chlorophyll derived from inversion of PROSAIL radiative transfer model on Sentinel-2A MSI image acquired on 21 July 2018. *A. dealbata and A. mearnsii* are shown in dark red colour.

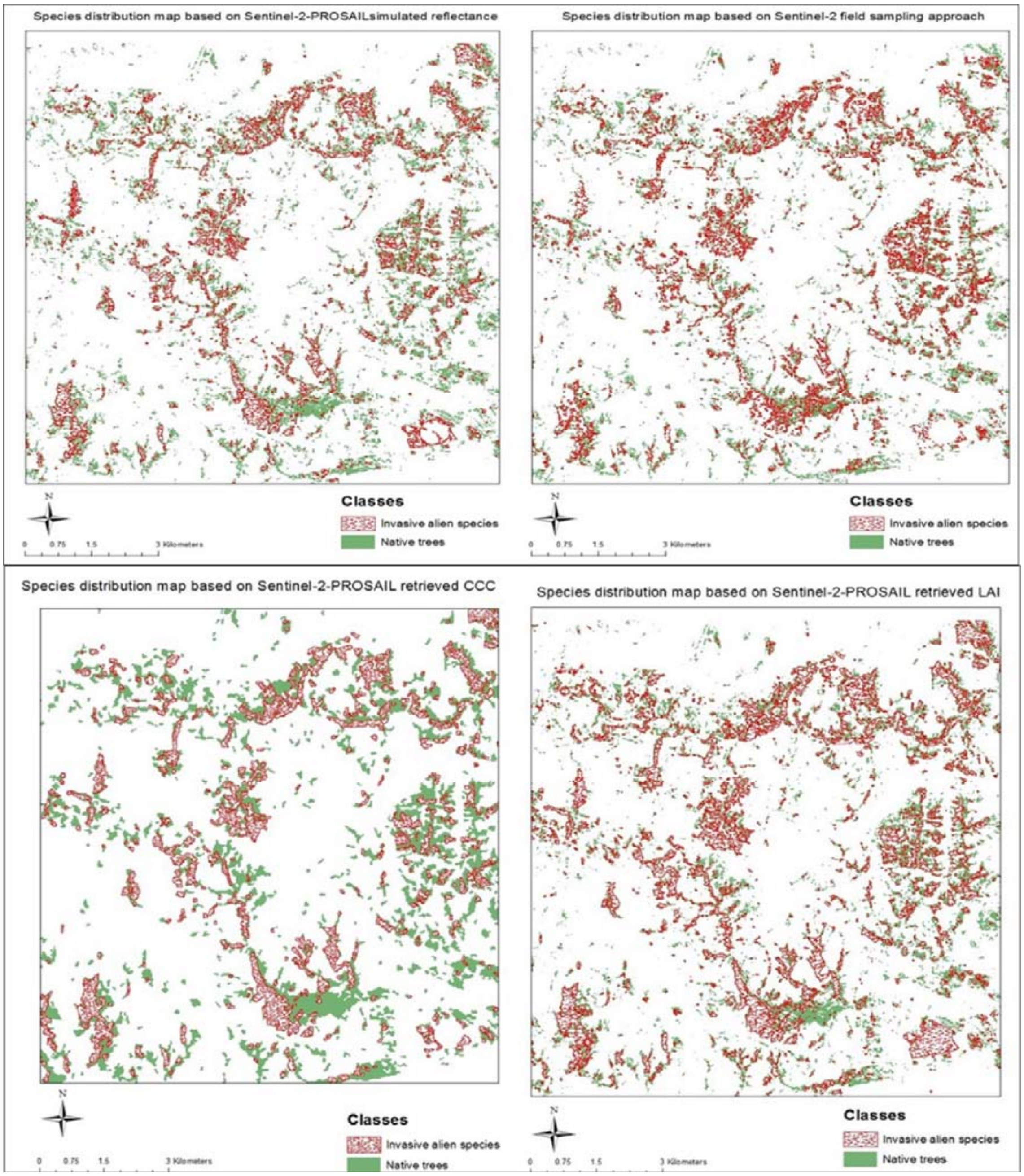


Fig. 9. Comparison between invasive alien plants and native species distribution maps based on field sampling data and equivalent maps derived from LAI and CCC maps obtained using PROSAIL inversion on Sentinel-2 MSI image acquired in December 2018. The dark red colour represents IAPs (*A. dealbata and A. mearnsii*).

5. Discussion

We have demonstrated that upscaling *in situ* leaf reflectance to canopy using canopy radiative transfer models can be used for *A. mearnsii* from native species at canopy scale. Furthermore, the study revealed that Sentinel-2 data, when combined with a radiative transfer model (PRO5SAIL4), has the potential to discriminate IAPs from native species at the landscape level and with satisfactory accuracy. Scaling-up leaf reflectance to canopy reflectance using the RT model produced promising results for the discrimination of *A. mearnsii* from native species. The synthetic canopy reflectance discriminated *A. mearnsii* with slightly lower accuracy and kappa coefficient when compared to measured canopy reflectance. In contrast, simulated spectral data showed less overlap between species than with measured species data (Fig. 4). According to Féret and Asner (2011) and (Clark, Roberts and Clark, 2005), misclassification of species is usually linked to the number of species. For example, Féret and Asner (2011) observed about 64% separability error when discriminating all 188 species. Conversely, when the species number was reduced to five classes, 90% discrimination accuracy was achieved, and ten categories decreased accuracy to 84%. However, in this study, we observed decreased spectral overlap between the species with synthetic data without reducing the number of species. Therefore, we attributed this to the ability of RTM to eliminate the effect of external atmospheric effects. Overall, these results supported (Kalacska et al., 2007, Cho et al., 2008, Féret and Asner, 2011) studies that successfully discriminated tree species by upscaling leaf reflectance to canopy reflectance.

This study showed that Sentinel-2 based PRO5SAIL5 models detected IAPs and native species with acceptable but slightly lower accuracies than the field sampling approach. The low performance of Sentinel-2 bands based PRO5SAIL5 model is related to the inability of the model to predict species in top hill areas. Fig. 9, demonstrate misclassification of species in some top hill areas. Most pixels were classified as “unclassified” pixels. Even though the RTM model showed low quantification statistics accuracy compared to the field-based model, the model still considered as promising because it was independent of field data. The model is considered promising because species distribution maps produced from PRO5SAIL4- Sentinel-2 MSI model exhibited similar patterns to that observed during the field survey. The same species patterns but high detection accuracies were observed with LAI and CCC retrieved with PRO5SAIL4 -Sentinel retrieval-based maps approach.

Overall, *A. mearnsii* was successfully discriminated from native species with higher accuracies when compared to *A. dealbata*. However, *A. mearnsii* and *A. dealbata* discrimination accuracies were comparable to that of field-based sampling method. Concerning LAI-based classification, both IAPs and native species indicated lower prediction accuracies, compared to the CCC-based mapping model.

The decreased accuracy of the LAI model may partly be due to underestimation of IAPs in mixed pixel situations and spatial resolution of sentinel-2 images. The mixed canopy pixels have been found to pose a challenge even in field-based mapping, particularly when IAPs canopies interlock with that of native species. We observed high accurate classification of IAPs and native species in monospecific stands than in mixed canopy stances, possibly due to inability of Random forest to un-mix the pixel. This proved the assumption that the PROSAIL model is useful in homogenous canopies (Darvishzadeh *et al.* 2008). Although PROSAIL Sentinel-2 MSI based model was adequate, the model depended mostly on the accurate parameterisation and inversion approach. For example, this study conducted a trial-and-error and species map based approach in which the range of values was tested against measured species Sentinel-2 reflectance.

The high performance of PROSAIL-Sentinel-2 based CCC retrieval model could be expained by availability of CCC specific bands of the image. The bands, for instance, red-edge bands have been reported to be highly sesnitive to the leaf chlorophyll concentration (Clevers and Gitelson, 2013). Furthermore, the red-edge bands are known to also sensitive to leaf nitrogen concentration (Ramoelo et al., 2015, Ramoelo and Cho, 2018) which is known to be high in *Acacia spp* (Große-Stoltenberg et al. 2018). The species have been reported to be nitrogen fixers, which makes them possess high leaf nitrogen contents when compared to non-fixing tree species. The importance of red-edge bands has also shown to be beneficial for PROSAIL Sentinel-2 MSI species mapping. The PRO5SAIL4-Sentinel-2 and DRF based classification found red-edge bands together with NIR and SWIR1 (1610 nm) and SWIR 2 (2190 nm) to be important predictors of the species. Furthermore, the bands are useful for distinguishing tree species based on differences in leaf tannin between species (Lehmann et al., 2015). This reinforces the fact that Australian native *Acacia* species have high leaf area and nitrogen/chlorophyll when compared to native species (Große-Stoltenberg et al., 2018, Masemola et al., 2019). Große-Stoltenberg et al. (2018) reported spectral dissimilarities between invasive acacias and indigenous plants in the biochemical space. Also, this validates Asner and Martin (2011), who demonstrated that biological traits mainly drive spectral separability of species. In Masemola *et al*. (under review), most of the spectra derivatives selected to be optimal for distinguishing *A. mearnsii* from native species were directly related to leaf physiology (leaf nitrogen and leaf chlorophyll). Importantly, the results showed that 20 m resolution Sentinel-2 bands are sufficient for the detection and mapping of IAPs. Therefore there is no need to resample-which could distort spectral information-when using simulated S2 configurations data for operational monitoring of investigated *Acacia* spp.

Comparing the RTMs (i.e., 1D and 3D RTMs) types of retrieval methods, it can be concluded that: (1) 1D RTM especially with it independence of field data is the most advantageous for operational mapping of IAPs. However, in terms of accuracy 1D RTM and field data based species mapping are comparable in terms of accuracy. However, RTMs seems to misclassify species at the rugged terrains. On the other hand, the 3D FLIGHT RTM slightly lower quantification statistics, compared to 1D RTM. This could be attributed to the fact that IAPs usually grow in clusters and form dense homogeneous canopy which meet simulation assumptions of 1D RTMs, that assume that everything is in turbid mode. Our results support that class-based inversion process may an approach of choice that can be used by land managers for automated monitoring of IAPs distribution. Unlike in Verrelst et al. (2012) study which used non-operational CHRIS data to explore class-based inversion approach, our results with Sentinel-2 data which is already being used for operational. This makes this study to be relevant to the current ecological problems related to biological invasion.

In general, the high performance of the developed PROSAIL-Sentinel-2 based model could play an essential role in *Acacia spp* monitoring. This is because, unlike field-based spectral statistical models, physical-based PROSAIL species modelling does not require larger calibration datasets before they can be implemented for operation methods. Furthermore, independence of PROSAIL Sentinel-2 MSI modelling to field species information and the fact that the models are transferable over a larger area indicate that they are suitable for operational mapping and monitoring purpose.

6. Conclusions

We have demonstrated that:

• Upscaling *in situ* leaf reflectance to canopy using canopy radiative transfer models can be used for *A. mearnsii* from native species at canopy scale.
• Sentinel 2A data, when combined with a radiative transfer model (PROSAIL), have the potential to discriminate IAPs from native species at the landscape level with satisfactory accuracy.
• Although PROSAIL Sentinel-2 MSI species detection was adequate, the model depended mostly on the accurate parameterisation.
• Leaf chlorophyll content and LAI derived from canopy radiative transfer models can be used to discriminate *A. mearnsii* at the larger scale.
• Sentinel-2 MSI bands are useful for mapping *A. mearnsii* at an ecosystem level.

Declaration of Competing Interest

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.

Acknowledgements

We thank the Council for Scientific and Industrial Research (CSIR), South Africa for proving the funding for the study. Lastly, we appreciate the very useful inputs of the anonymous reviewers.

References

Asner, G.P., Martin, R.E. Spectranomics: emerging science and conservation opportunities at the interface of biodiversity and remote sensing. Glob. Ecol. Conserv., 8 (2016), pp. 212-219

Baret, F., Guyot, G. Potentials and limits of vegetation indices for LAI and APAR assessment. Remote Sens. Environ., 35 (2–3) (1991), pp. 161-173, 10.1016/0034-4257(91)90009-U

Berk, A., Anderson, G., … P.A.-A. and, 2006, undefined, n.d. MODTRAN5: 2006 update. spiedigitallibrary.org.

Chamier, J., et al. Impacts of invasive alien plants on water quality, with particular emphasis on South Africa. Water SA, 38 (2) (2012), pp. 345-356

Chance, C.M., et al. Invasive shrub mapping in an urban environment from hyperspectral and LiDAR-derived attributes. Front. Plant Sci., 7 (2016), p. 1528

Cho, M.A., et al. Discriminating species using hyperspectral indices at leaf and canopy scales. Int. Arch. Photogr. Remote Sensing Spatial Information Sci., 37 (B7) (2008)

Clark, M.L., Roberts, D.A., Clark, D.B. Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. Remote Sens. Environ., 96 (3–4) (2005), pp. 375-398

Clevers, J.G.P.W., Gitelson, A.A. Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. Remote Sens. Environ., 96 (3–4) (2005), pp. 375-398

Combal, B., Baret, F., Weiss, M., Trubuil, A., Macé, D., Pragnè, A., Myneni, R., Knyazikhin, Y., Wang, L. Retrieval of canopy biophysical variables from bidirectional reflectance Using prior information to solve the ill-posed inverse problem. (2002)

Delegido, J., Verrelst, J., Rivera, J.P.,Ruiz-Verdú, A., Moreno, J. Brown and green LAI mapping through spectral indices. Int. J. Appl. Earth Observat. Geo-information, 35 (2015), pp. 350-358

Dorigo, W. Retrieving canopy variables by radiative transfer model inversion A regional approach for imaging spectrometer data. (2008) Available at: https://www.geo.tuwien.ac.at/downloads/wd/thesis/dorigo2008_phdthesis.pdf (Accessed: 17 April 2019).

Féret, J.B., François, C., Asner, G. PROSPECT-4 and 5: Advances in the leaf optical properties model separating photosynthetic pigments. Remote sensing of environment, 112(6) (2008), 3030-3043.

Fu, H., Zhang, Q., Qiu, G. Random forest for image annotation. In: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) (2012) pp. 86–99. https://doi-org.uplib.idm.oclc.org/10.1007/978-3-642-33783-3_7.

Ge, H., Lu, D., He, S., Xu, A., G.Z.-...E.R. undefined,2008.06–085.qxd,ingentaconnect.com. (2008)

Goel, N.S. Models of vegetation canopy reflectance and their use in estimation of biophysical parameters from reflectance data. Remote Sensing Reviews, 4 (1) (1988), pp. 1-212, 10.1080/02757258809532105

Große-Stoltenberg, A., et al. Evaluation of continuous VNIR-SWIR spectra versus narrowband hyperspectral indices to discriminate the invasive Acacia longifolia within a Mediterranean dune ecosystem. Remote Sensing, 8 (4) (2016), p. 334

Große-Stoltenberg, A., et al. Invasive acacias differ from native dune species in the hyperspectral/biochemical trait space. J. Veg. Sci., 29 (2) (2018), pp. 325-335

Großkinsky, D.K., et al. Plant phenomics and the need for physiological phenotyping across scales to narrow the genotype-to-phenotype knowledge gap. J. Exp. Bot., 66 (18) (2015), pp. 5429-5440

Hellmann, C., et al. Heterogeneous environments shape invader impacts: integrating environmental, structural and functional effects by isoscapes and remote sensing. Sci. Rep., 7 (1) (2017), p. 4118

Impson, F.A.C., Kleinjan, C.A., Hoffmann, J.H., Post, J.A. Dasineura rubiformis (Diptera: Cecidomyiidae), a new biological control agent for Acacia mearnsii in South Africa. South African J. Sci., 104 (2008), pp. 247-249

Jacquemoud, S., Baret, F. PROSPECT: A model of leaf optical properties spectra. Remote Sens. Environ., 34 (2) (1990), pp. 75-91

Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P.J., Asner, G.P., François, C., Ustin, S.L. PROSPECT + SAIL models: a review of use for vegetation characterization. Rem. Sens. Environ., 113 (2009), pp. S56-S66, 10.1016/j.rse.2008.01.026

Jiménez, M.,Díaz-Delgado, R. Towards a standard plant species spectral library protocol for vegetation mapping: A case study in the shrubland of Doñana National Park. ISPRS Int. J. Geo-Inf., 4 (4) (2015), pp. 2472-2495

Kalacska, M., et al. Hyperspectral discrimination of tropical dry forest lianas and trees: Comparative data reduction approaches at the leaf and canopy levels. Remote Sens. Environ., 109 (4) (2007), pp. 406-415, 10.1016/j.rse.2007.01.012

Lázaro-Gredilla, M., Titsias, M.K., Verrelst, J., Camps-Valls, G. Retrieval of biophysical parameters with heteroscedastic Gaussian processes. IEEE Geosci. Remote Sens. Lett., 11 (4) (2014), pp. 838-842

Lehmann, J.R.K., et al. Field spectroscopy in the VNIR-SWIR region to discriminate between Mediterranean native plants and exotic-invasive shrubs based on leaf tannin content. Remote Sensing, 7 (2) (2015), pp. 1225-1241

LeMaitre, D.C., et al. Impacts of invasive Australian acacias: implications for management and restoration. Divers. Distrib., 17 (5) (2011), pp. 1015-1029

Makkonen, M. et al. 'Highly consistent effects of plant litter identity and functional traits on decomposition across a latitudinal gradient', Ecology Letters. Edited by J. Klironomos, 15(9) (2012), pp. 1033–1041. doi: 10.1111/j.1461-0248.2012.01826.x.

Masemola, C., Cho, M.A., Ramoelo, A. Assessing the effect of seasonality on leaf and canopy spectra for the discrimination of an alien tree species, acacia mearnsii, from co-occurring native species using parametric and nonparametric classifiers. IEEE Trans. Geosci. Remote Sens., 57 (8) (2019), 10.1109/TGRS.2019.2902774

Morris, T.L., et al. Ecophysiological traits associated with the competitive ability of invasive Australian acacias. Divers. Distrib., 17 (5) (2011), pp. 898-910

Naidoo, L., Cho, M.A., Mathieu, R., Asner,G. , Classification of savanna tree species, in the Greater Kruger National Park region, by integrating hyperspectral and LiDAR data in a Random Forest data mining environment. ISPRS J. Photogram. Rem. Sens., 69 (2012), pp. 167-179

Pandey, P.C., Tate, N.J., Balzter,H. Mapping tree species in coastal portugal using statistically segmented principal component analysis and other methods. IEEE Sens. J., 14 (12) (2014), pp. 4434-4441, 10.1109/JSEN.2014.2335612

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V. Scikit-learn: machine learning in Python. J. Mach. Learn. Res., 12 (2011), pp. 2825-2830

Punalekar ,S.M., et al. Application of Sentinel-2A data for pasture biomass monitoring using a physically based radiative transfer model. Remote Sensing Environment. Elsevier, 218 (October) (2018), pp. 207-220, 10.1016/j.rse.2018.09.028

Ramoelo, A., et al. Potential of Sentinel-2 spectral configuration to assess rangeland quality. J. Appl. Remote Sens., 9 (1) (2015), p. 94096

Ramoelo, A., Cho, M.A. Explaining leaf nitrogen distribution in a semi-arid environment predicted on sentinel-2 imagery using a field spectroscopy derived model. Remote Sensing, 10 (2) (2018), p. 269

Rascher, K.G., et al. Community scale 15N isoscapes: tracing the spatial impact of an exotic N2-fixing invader. Ecol. Lett., 15 (5) (2012), pp. 484-491

Rivera, J.P., Verrelst, J., Leonenko, G., Moreno, J. Multiple cost functions and regularization options for improved retrieval of leaf chlorophyll content and LAI through inversion of the PROSAIL model. Rem. Sens., 5 (2013), pp. 3280-3304

Rivera-Caicedo, J.P., et al. Hyperspectral dimensionality reduction for biophysical variable statistical retrieval. ISPRS J. Photogramm. Remote Sens., 132 (2017), pp. 88-101, 10.1016/j.isprsjprs.2017.08.012

Savitzky, A., Golay, M.J.E. Smoothing and differentiation of data by simplified least squares procedures. Anal. Chem., 36 (1964), pp. 1627-1639

Simberloff, D., Martin, J.L., Genovesi, P., Maris, V., Wardle, D.A., Aronson, J., Pyšek,P. Impacts of biological invasions: what's what and the way forward. Trends Ecol. Evol., 28 (1) (2013), pp. 58-66

Skowronek, S., et al. Mapping an invasive bryophyte species using hyperspectral remote sensing data. Biol. Invasions, 19 (1) (2017), pp. 239-254

Sun, J., Shi, S., Yang, J., Du, L., Gong, W., Chen, B.W., Song, S.L. Analyzing the performance of PROSPECT model inversion based on different spectral information for leaf biochemical properties retrieval. ISPRS J. Photogramm., 135 (2018), pp. 78-83

Thenkabail, P.S. Land resources monitoring, modeling, and mapping with remote sensing. CRC Press (2015)

Thenkabail, P.S., Lyon, J.G. Hyperspectral remote sensing of vegetation. CRC Press (2016)

Thenkabail, P.S., Lyon, J.G., Huete, A Hyperspectral remote sensing of vegetation. Volume III, Biophysical and biochemical characterization and plant species stud. (2019)

Van Wittenberghe, S., et al. Gaussian processes retrieval of leaf parameters from a multi-species reflectance, absorbance and fluorescence dataset. J. Photochem. Photobiol. B: Biol. Elsevier, 134 (2014), pp. 37-48, 10.1016/J.JPHOTOBIOL.2014.03.010

Van Wilgen, B.W. and Wilson, J.R. 'The status of biological invasions and their management in South Africa in 2017', South African National Biodiversity Institute, Kirstenbosch and DST-NRF Centre of Excellence for Invasion Biology, Stellenbosch, (2018), p. 398.

Verhoef, W. Light scattering by leaf layers with application to canopy reflectance modeling: the SAIL model. Remote Sens. Environ., 16 (1984), pp. 125-141

Verrelst, J., Muñoz, J., Alonso, L., Delegido, J., Rivera, J.P., Camps-Valls, G., Moreno, J. Machine learning regression algorithms for biophysical parameter retrieval: Opportunities for Sentinel-2 and-3. Remote Sens. Environ., 118 (2012), pp. 127-139

Verrelst, J., Camps-Valls, G., Muñoz-Marí, J., Rivera, J.P., Veroustraete, F., Clevers, J.G.P.W., Moreno, J. Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties – A review. ISPRS J. Photogramm. Remote Sens., 108 (2015), pp. 273-290

Verrelst, J., Rivera, J.P., Moreno, J., Camps-Valls, G. Gaussian processes uncertainty estimates in experimental Sentinel-2 LAI and leaf chlorophyll content retrieval. ISPRS J. Photogramm. Remote Sens., 86 (2013), pp. 157-167

Verrelst, J., Rivera, J.P., Veroustraete, F., Munoz-Mari, J., Clevers, J.G.P.W.,Camps-Valls, G., Moreno, J. Experimental Sentinel-2 LAI estimation using parametric, nonparametric and physical retrieval methods – A comparison. ISPRS J. Photogramm., 108 (2015), pp. 260-272

Xu, X.Q., Lu, J.S., Zhang, N., Yang, T.C., He, J.Y., Yao, X., Cheng, T., Zhu, Y., Cao, W.X., Tian, Y.C Inversion of rice canopy chlorophyll content and leaf area index based on coupling of radiative transfer and Bayesian network models. ISPRS J. Photogramm. Remote Sens., 150 (2019), pp. 185-196