



Testing the efficacy of hyperspectral (AVIRIS-NG), multispectral (Sentinel-2) and radar (Sentinel-1) remote sensing images to detect native and invasive non-native trees

M. Arasumani · Aditya Singh · Milind Bunyan · V. V. Robin

Received: 17 December 2020 / Accepted: 16 April 2021 / Published online: 27 April 2021
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2021

Abstract Invasive alien species threaten tropical grasslands and native biodiversity across the globe, including in the natural mosaic of native grasslands and forests in the Shola Sky Islands of the Western Ghats. Here, grasslands have been lost to exotic tree invasion (Acacias, Eucalyptus, and Pines) since the 1950s, but differing invasion intensities between these species and intermixing with native species constitutes a major challenge for remotely sensed assessments. In this study, we assess the accuracy of three satellite and airborne remote sensing sensors (Sentinel-1 radar data, Sentinel-2 multispectral data and AVIRIS-NG hyperspectral data) and three machine learning

classification algorithms to identify the spatial extent of native habitats and invasive tree species. We used the support vector machine (SVM), classification and regression trees (CART), and random forest (RF) algorithms implemented on the Google Earth Engine platform. Results indicate that AVIRIS-NG data in combination with SVM produced the highest classification accuracy (98.7%). Fused Sentinel-1 and Sentinel-2 produce 91% accuracy, while Sentinel-2 alone yielded 91% accuracy; but only with higher coverage of ground control points. The hyperspectral data (AVIRIS-NG) was the only sensor that permitted distinguishing recent invasions (young trees) with high precision. We suspect that large areas will have to be mapped and assessed in the coming years by conservation managers, NGOs to plan restoration or to assess the success of restoration activities, for which a choice of sensors may have to be made based on the age of invasion being mapped, and the quantum of ground control data available.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10530-021-02543-2>.

M. Arasumani (✉) · V. V. Robin
Department of Biology, Indian Institute of Science
Education and Research (IISER) Tirupati, Transit
Campus, Karakambadi Road, Tirupati,
Andhra Pradesh 517507, India
e-mail: arasustat@gmail.com

A. Singh
Department of Agricultural and Biological Engineering,
University of Florida, 1741 Museum Rd., Gainesville,
FL 32611, USA

M. Bunyan
Ashoka Trust for Research in Ecology and the
Environment (ATREE), Royal Enclave, Srirampura,
Bangalore, Karnataka 560064, India

Keywords AVIRIS-NG · Sentinel-1 · Sentinel-2 ·
Google Earth Engine · Invasive species mapping ·
Western Ghats · Tropical montane grasslands

Introduction

Invasive alien tree species threaten ecosystem integrity by modifying the structure and function of ecosystems and have negative impacts on ecosystem services delivery and native biodiversity (Le Maitre et al. 2011; Mooney 2005). In particular, three genera of invasive alien trees—acacias, pines and eucalyptus species are often listed as the worst offenders in South Africa (Gaertner et al. 2017; McConnachie et al. 2015), Brazil (de Abreu and Durigan 2011), Argentina (Zalba et al. 2008), Uruguay (Six et al. 2013), Kenya (Pellikka et al. 2009), New Zealand (Ledgard 2001), Hawaiian Islands (Daehler 2005) and India (Arasumani et al. 2019). These trees are native to Australia and were introduced to several tropical and subtropical countries in the nineteenth century (Richardson 1998). These trees were largely established on grasslands and shrublands which were historically classified as wastelands (Joshi et al. 2018; Rundel et al. 2014). All of these species are fast-growing, have high water use and show potential for encroachment into native grasslands and scrublands. Acacias, in particular form dense stands, maintain a high leaf area all through the year and have high evapotranspiration which impacts water yields of infested catchments (Dye and Jarman 2004). In South Africa, these species have also had negative impacts on grassland birds (Allan et al. 1997; Armstrong and Van Hensbergen 1995), small mammals (Armstrong and Van Hensbergen 1995), invertebrates (Donnelly 1985) and plants (Richardson and Van Wilgen 1986).

In the Western Ghats, tropical montane grasslands, also known as Shola grasslands have been lost to exotic tree invasion at a rapid pace (Arasumani et al. 2018, 2019). This widespread invasion has impacted various faunal communities, including threatened species like the Nilgiri Pipit (Lele et al. 2020) and Nilgiri Tahr (Alempath 2008) in the Western Ghats. These trees were primarily established outside wildlife preserves and protected areas since the 1950s. However, today, the invasion of exotic trees is rapidly encroaching critical grassland landscapes, including protected areas (Arasumani et al. 2019; Joshi et al. 2018). Our prior research using Landsat data (Arasumani et al. 2019) indicates that 23% (340 sq. km.) of montane grasslands have been lost to invasive exotic tree stands within the past five decades. This data, however, had included all exotic trees as a single class

due to the limitations with the spectral and spatial resolution of LANDSAT imagery. Maps classifying multiple invasive species into a single-class have limitations as some species, like Acacia and Pine, may invade remnant grasslands more rapidly than others. In the Western Ghats, *Acacia mearnsii* and *Pinus patula*, both listed among the most invasive species in the World (IUCN-GISD), co-occur with less invasive timber species as well as with native tropical forests (Joshi et al. 2018). Such a complex landscape makes the detection and estimation of the spread of the invasive trees challenging. One of the major challenges for conservation managers is to detect the invasion front, which typically consists of sparsely dispersed seedlings and saplings that are difficult to detect using remote sensing imageries (but see Arasumani et al. 2020). Detecting this invasion front is, however, critical in managing the invasion front as is identifying the best combination of algorithms and remote sensing platforms that can be used by conservation agencies in tropical areas across the globe.

Choice of imageries

Field-based, landscape-scale, invasive species mapping techniques are known to be challenging in tropical forests. Many researchers have attempted to identify broad forest type categories using space-borne and airborne images (Arasumani et al. 2018; Erinjery et al. 2018; Foody and Hill 1996; Shimizu et al. 2019). Discriminating between native and non-native species, however, remains challenging with medium resolution satellite data due to extensive intermixing of species and canopy heterogeneity. Such complex landscape scenarios may require high spatial and spectral resolution imageries for appropriately classifying invasive species. Researchers have also employed microwave data (Chen et al. 2018; Laurin et al. 2013; Wheeler et al. 2017), recently launched Sentinel-2 multispectral images (Laurin et al. 2013; Wheeler et al. 2017), and combinations of Sentinel-1 SAR (Synthetic Aperture Radar) images with Sentinel-2 multispectral images to improve the classification accuracy of tropical forests types and woody invasive species (Erinjery et al. 2018; Kattenborn et al. 2019; Zhang et al. 2019). A few studies have also suggested that hyperspectral images could produce high accuracy for mapping tree species compared to

the multispectral datasets (Awad 2018; George et al. 2014; Thenkabail et al. 2004).

Choice of classification algorithms

Several classification algorithms have been used for classifying hyperspectral, multispectral, and SAR data: of these, support vector machine (SVM; Burgess et al. 2007), random forest (RF; Breiman 2001), and classification and regression tree (CART; Breiman et al. 1984) have been the most widely used especially on the publicly-available Google Earth Engine Platform (Gorelick et al. 2017). Although some studies pick the best classification algorithm for classifying hyperspectral data; all three algorithms are used for mapping landcover on both multispectral and SAR images (Lu et al. 2018) with perhaps differences in their training data requirements. The SVM is an iterative, non-parametric machine learning algorithm widely used for classifying hyperspectral images (Mountrakis et al. 2011). The SVM algorithm does not depend on the statistical distribution of the data but relies on training data adjacent to the class boundary to deliver high accuracy even with limited training data for classification (Melgani and Bruzzone 2004). The random forest (RF) algorithm is a non-parametric classifier that operates by generating a number of classification trees and selecting the mode of the predictions (Breiman 2001). CART models use recursive binary splits on predictor data in a decision tree framework to produce classifications at the end nodes of the trees. By nature of the classification process, CART models are considered somewhat easier to interpret compared to RF-based models (Lawrence and Wright 2001). Notably, RF and CART approaches are widely used to classify the remote sensing images but have higher training data requirements relative to SVM models (Delalay et al. 2019; Shaharum et al. 2020).

Objectives

Overall, this study aims to identify appropriate data sources and algorithms to identify exotic tree species on the Shola Sky Islands of the Western Ghats. Specifically, we test the ability of (a) AVIRIS-NG (hyperspectral), Sentinel-2B (multispectral) and Sentinel-1B (microwave) data with (b) SVM, Random Forest, and CART classification algorithms to

discriminate invasive woody species (*Acacia*, Pine and *Eucalyptus*) from native tropical trees in the Shola Sky Islands. The overall goal is to test the efficacy of these imageries in detecting the invasion front—the ecotone where conservation efforts can be targeted.

Methods

Study area

For the purposes of this study, we selected an area admeasuring approximately 12 sq. km. in the Nilgiris mountains (Fig. 1) that has a gradient of woody invasive species infestation across this landscape. The study area contains natural habitats such as montane grasslands, montane forests and water bodies, and non-native woody trees—*Acacia* spp, *Eucalyptus* spp, and *Pinus* spp (Fig. 2).

Field data

In our study region, native forests are generally intermixed with non-native trees (*Acacia* spp., *Pinus* spp., and *Eucalyptus* spp.). However, only acacias were observed to be present at the actual grassland-forest invasion front (i.e., pines and eucalyptus were generally established in distinct stands). We collected the unique land cover GPS points using simple random sampling where target species land covered in excess of a 30 m*30 m footprint to obtain ‘pure’ endmember spectral signatures. We used 1049 ground-truth points of montane grasslands, forest, acacias, eucalyptus, pines and water bodies for image classification. For all locations, we ensured that the positional dilution of precision (PDOP) was lower than the AVIRIS-NG, Sentinel-1, and Sentinel-2- spatial resolutions (4 m and 10 m). To detect the *Acacia* spp. invasion front, we collected 73 additional ground truth locations concurrent with satellite imagery. We assessed the efficiency of acacia invasion front mapping with classifications produced by the three different image datasets and algorithms on the Google Earth Engine Platform as described below.

Image data

We obtained an AVIRIS-NG (Airborne Visible InfraRed Imaging Spectrometer—Next Generation)

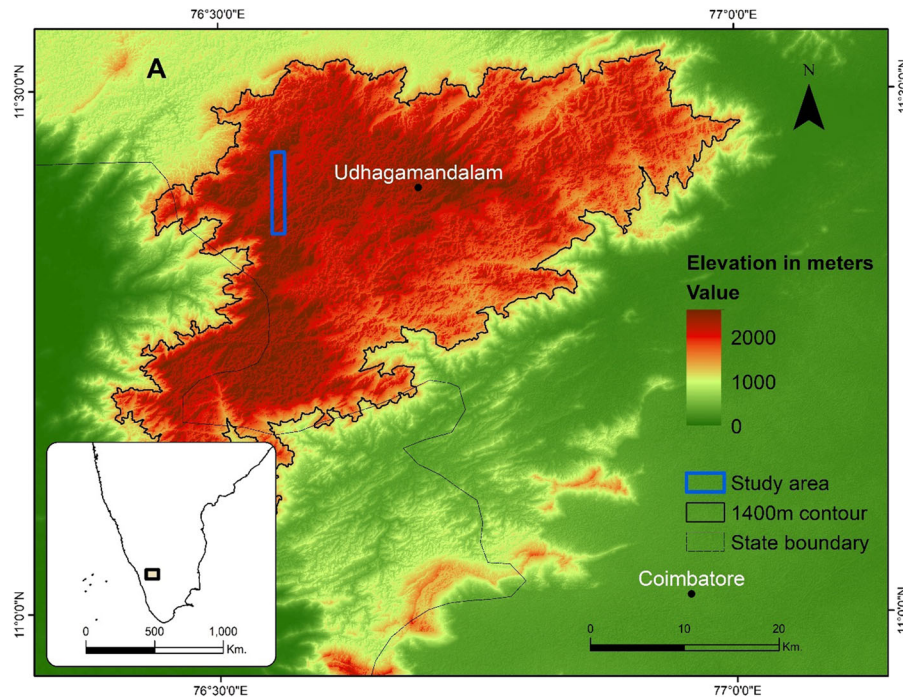


Fig. 1 a Study area—Nilgiri Hills. b A—AVIRIS-NG false colour composite image, B—Sentinel—2 false colour composite image, C—Sentinel—1 VV polarisation image with training data and D—Sentinel—2 VH polarisation with validation data

apparent at-surface reflectance products (March 2018) from the Jet Propulsion Laboratory (JPL), National Aeronautics and Space Administration data portal (NASA—<https://avirisng.jpl.nasa.gov/dataportal/>). AVIRIS-NG data have high spatial (4 m) and spectral (5 nm) spectral resolution with 425 spectral bands spanning 380–2510 nm. We excluded noisy bands and water vapour absorption bands (bands 1–10, 195–207, 287–316, 325–329) from the AVIRIS-NG dataset.

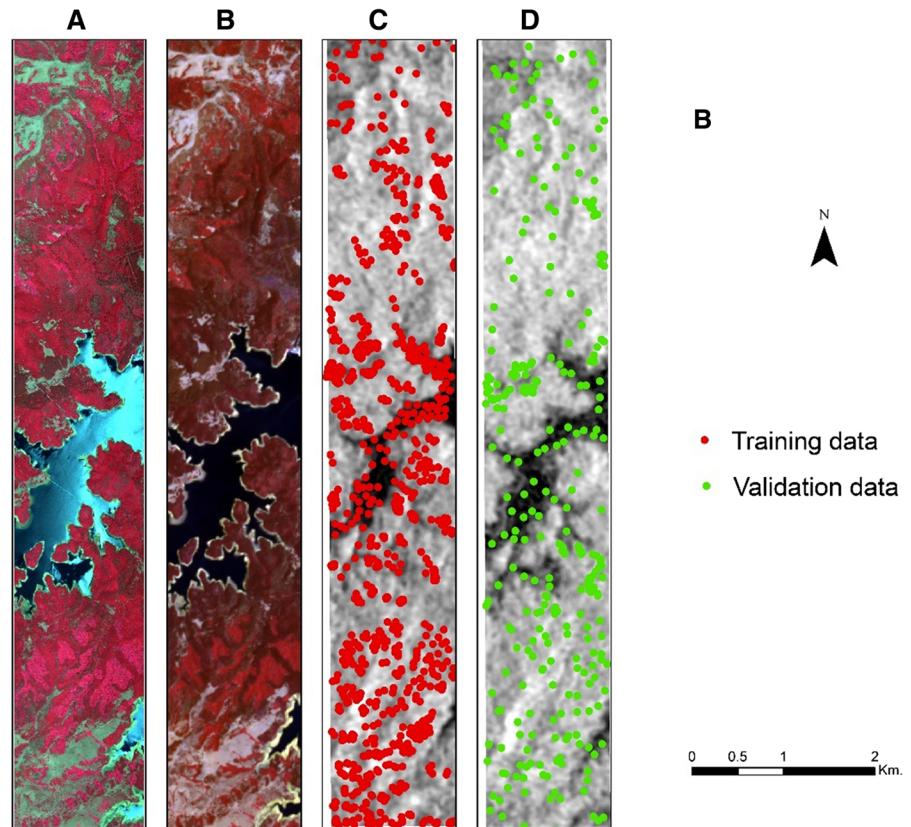
The Sentinel-2 Level-2 ground reflectance product was not available for our study area on the GEE platform for 2018. We, therefore, substituted this with the Sentinel-2 Level-1 product from the USGS (United States Geological Survey) Earth Explorer portal (<https://earthexplorer.usgs.gov/>), and we performed atmospheric corrections using Sen2Cor v2.8 in the Sentinel Application Platform (SNAP) to convert at-sensor radiance imagery to apparent at-surface reflectance (http://step.esa.int/main/snap-supported-plugins/sen2cor/sen2cor_v2-8/). We standardized all bands to a 10 m spatial resolution for all subsequent analyses.

We obtained the Sentinel-1 SAR Ground Range Detected (GRD) product (March 2018) directly from the GEE platform. This data was available as calibrated and ortho-corrected and was pre-processed using the Sentinel-1 toolbox for thermal noise removal, radiometric calibration, and terrain correction (using the SRTM 30 m spatial resolution digital elevation model). The final terrain-corrected data were log-transformed to decibels. We used Sentinel-1 VH (vertical transmit and horizontal receive) and VV (vertical transmit and vertical receive) polarization for image classification. We fused the Sentinel-1 (VV, VH) bands (spatial resolution 10 m) and Sentinel-2 (2, 3, 4, 5, 6, 7 and 8) bands (spatial resolution 10 m and 20 m) for all subsequent analyses.

Classification algorithms

All image classifications were conducted using the Google Earth Engine (GEE) platform. We chose GEE to enable the creation of a workflow that can be utilized by conservation managers in other regions to track and manage invasions. We tested a combination

Fig. 1 continued



of Random Forest (RF; Breiman 2001), Classification and Regression Tree (CART; Breiman et al. 1984), and Support Vector Machine using radial basis functions (SVM; Burges 1998) to assess the skill of classification. We iterated the random forest trees from 0 to 50. We find that the RF classification accuracy improves as the number of trees increases; however, this pattern is seen in up to 6 trees. The classifier's performance continues almost identical to 50 trees. For the SVM, parameter values (gamma and cost) were determined by using an iterative grid search; gamma 0.6 and cost 220 produced the high classification accuracy.

We assessed the accuracy of the classifications using an additional set of 300 ground truth points held out from all preceding analyses and estimated accuracy metrics (overall accuracy, user accuracy, producer accuracy and, Kappa coefficient) from the confusion matrix (Congalton 1991; Congalton and Green 2019).

Results

Comparison of different classification

We observed the highest overall accuracy (OA; 98.7%) and kappa coefficient (Kappa; 0.984) in AVIRIS-NG dataset with SVM classification (Table 1, Fig. 3), followed by S1 + S2 datasets with CART (OA—91.3%, Kappa—0.896), Sentinel-2 with SVM (OA—91.0%, Kappa—0.892) and Sentinel-2 with RF (OA—91.0%, Kappa—0.892).

The highest overall accuracy in the AVIRIS-NG dataset was recorded with SVM classification (OA—98.7, Kappa—0.984), followed by Random Forest classification (OA—88.7 Kappa—0.864) and CART classification (OA—85.7 Kappa—0.828). (Table 1, Fig. 4).

The highest overall accuracy for Sentinel-2 data was also obtained with SVM classification (OA—91.0, Kappa—0.892) and RF classification (OA—91.0, Kappa—0.892), followed by CART

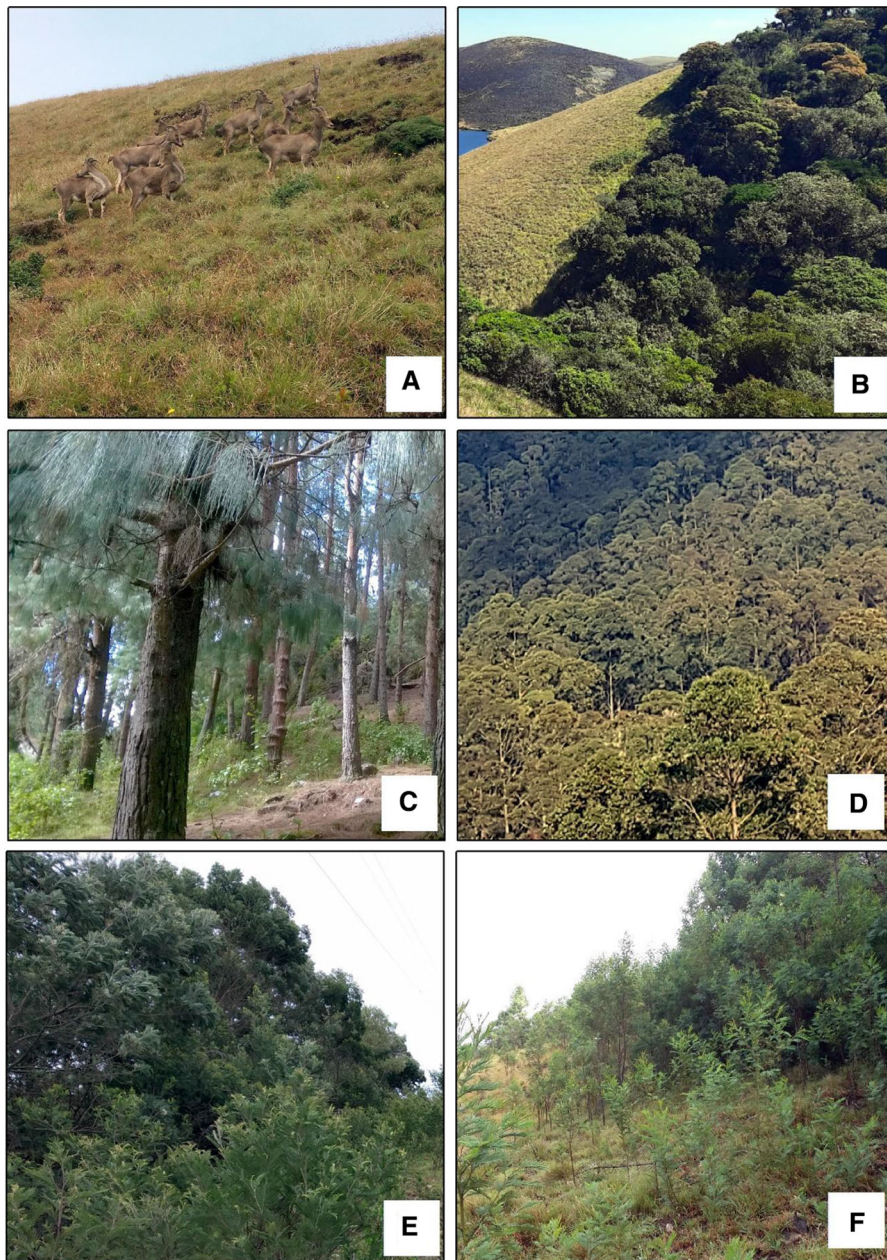


Fig. 2 Field photographs; **a** Montane grasslands, **b** Montane forests, **c** Pine, **d** Eucalyptus, **e** Acacia, and **f** Acacia invasion front

classification (OA—86.3, Kappa—0.836) with S2 bands of 2, 3, 4, 5, 6, 7 and 8 (Table 1, Fig. 5).

The S1 + S2 datasets provided the highest overall accuracy with CART classification (OA—91.3, Kappa—0.896) and RF classification (OA—90.0, Kappa—0.88), followed by SVM classification (OA—82.0, Kappa—0.784) with Sentinel-2 bands 2,

3, 4, 5, 6, 7 and 8 and Sentinel-1 VV and VH polarisations (Table 1, Fig. 6).

The Sentinel-1 dataset resulted in low accuracy of all classifiers compared to the other datasets with the highest overall accuracy with RF classification (OA—49.3, Kappa—0.392), and CART classification (OA—48.0, Kappa—0.376), followed by SVM classification

(OA—48.3, Kappa-0.38) with VV and VH polarisations (Table 1, Fig. 7).

Best producer and user accuracy

The highest producer accuracy and user accuracy were obtained from the AVIRIS-NG dataset with SVM of all classes. Both Sentinel-2 imageries with SVM classification, and S1 + S2 datasets with CART classification produced high producer & user accuracy in detecting water, forests and grasslands, but had lower accuracy in identifying acacias, eucalyptus and pine. However, Sentinel-1 performed poorly, and the user and producer accuracy of all the classes was low (< 50%) except for water (Table 1).

Number of training samples vs classification accuracy

With limited data (25% of training samples), the AVIRIS-NG dataset with SVM classification

performed well (OA—94%, Kappa—0.93) but RF and CART classifications yielded a lower OA (81%; Table 2). The Sentinel-2 with RF classification performed modestly (OA—89%, Kappa—0.85) with 25% of training samples. We observed that AVIRIS-NG dataset with SVM classification produced very high OA (97%) even with the 50% per cent of training samples; however, Sentinel-2 and S1 + S2 datasets yielded an overall accuracy of less than 90%. We noted that the classification accuracy for S1 + S2 datasets did not vary much with partial training datasets of 75% when compared to the complete training data set (Table 2).

Detecting the acacia invasion front

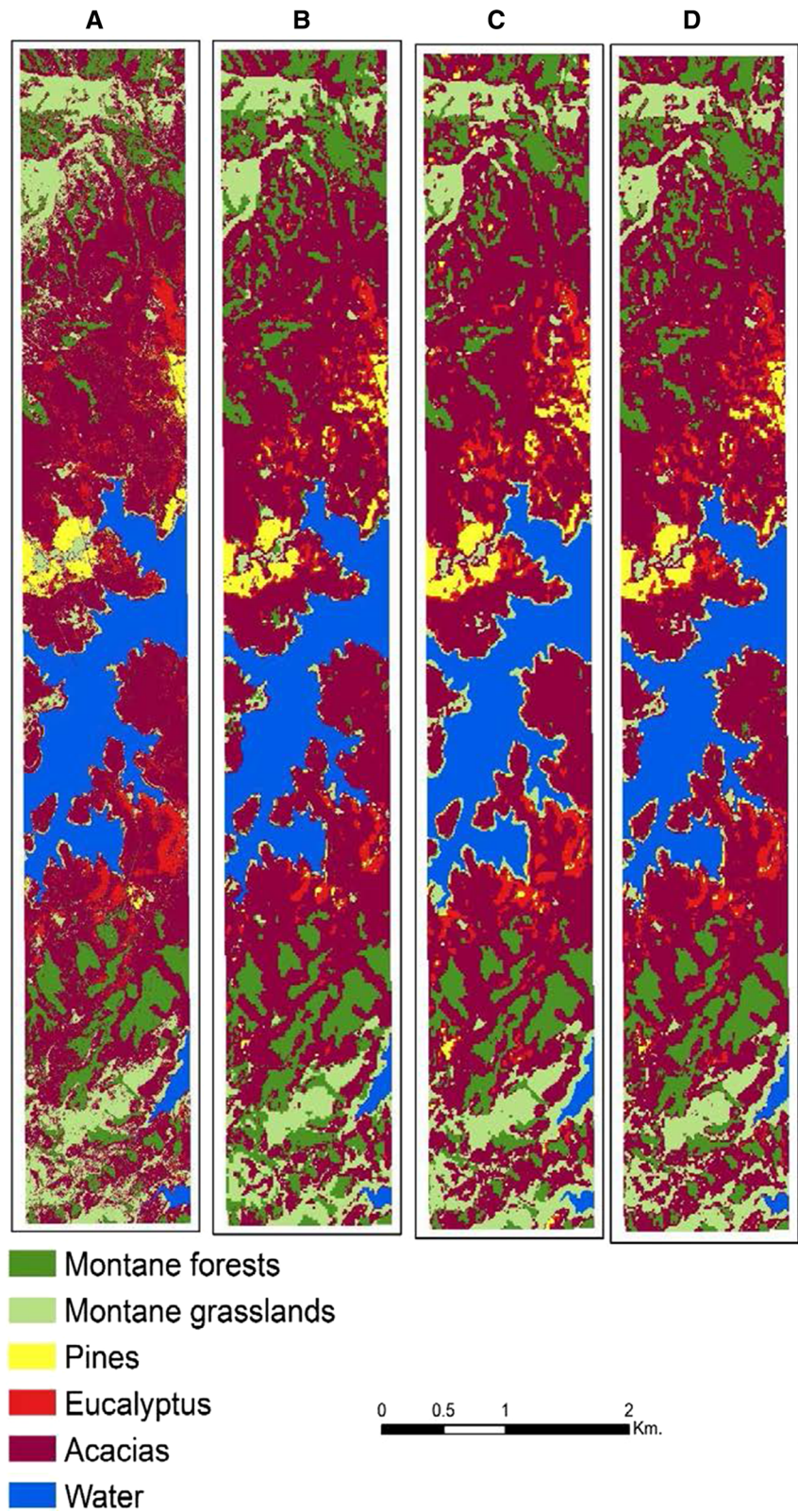
We observed high accuracy in detecting acacia invasion front with AVIRIS-NG dataset with SVM classification (97%). The RF and CART classification, however, produced lower accuracy (< 40%). Moderate accuracy was observed in the Sentinel-2 with RF

Table 1 Image classification results

| Landcover class | AVIRIS (SVM) | | AVIRIS (RF) | | AVIRIS (CART) | | S2 (SVM) | | S2 (RF) | | S2 (CART) | |
|--------------------|--------------|-------|-------------|-------|---------------|-------|----------|-------|---------|-------|-----------|-------|
| | UA | PA | UA | PA | UA | PA | UA | PA | UA | PA | UA | PA |
| Montane forest | 100.0 | 94.0 | 100.0 | 96.0 | 100.0 | 90.0 | 98.0 | 100.0 | 96.1 | 98.0 | 89.3 | 100.0 |
| Montane grasslands | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 98.0 | 100.0 | 98.0 | 100.0 | 98.0 | 100.0 | 90.0 |
| Acacias | 98.0 | 98.0 | 86.0 | 86.0 | 74.1 | 86.0 | 88.0 | 88.0 | 78.0 | 92.0 | 69.7 | 92.0 |
| Pines | 100.0 | 100.0 | 70.8 | 92.0 | 71.0 | 88.0 | 79.6 | 86.0 | 85.7 | 84.0 | 89.3 | 78.0 |
| Eucalyptus | 98.0 | 100.0 | 78.4 | 58.0 | 76.5 | 52.0 | 80.4 | 74.0 | 88.1 | 74.0 | 81.1 | 60.0 |
| Water | 100.0 | 100.0 | 100.0 | 100.0 | 96.2 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 98.0 |
| Overall accuracy | 98.7 | | 88.7 | | 85.7 | | 91.0 | | 91.0 | | 86.3 | |
| Kappa coefficient | 0.984 | | 0.864 | | 0.828 | | 0.892 | | 0.892 | | 0.836 | |
| Landcover class | S1&2 (SVM) | | S1&2 (RF) | | S1&2 (CART) | | S1 (SVM) | | S1 (RF) | | S1 (CART) | |
| | UA | PA | UA | PA | UA | PA | UA | PA | UA | PA | UA | PA |
| Montane forest | 90.6 | 96.0 | 90.9 | 100.0 | 98.0 | 100.0 | 33.3 | 38.0 | 31.1 | 38.0 | 28.6 | 20.0 |
| Montane grasslands | 88.6 | 78.0 | 100.0 | 92.0 | 100.0 | 98.0 | 75.9 | 44.0 | 75.0 | 54.0 | 67.6 | 46.0 |
| Acacias | 76.2 | 90.0 | 79.3 | 92.0 | 78.3 | 94.0 | 28.0 | 52.0 | 24.7 | 48.0 | 26.6 | 58.0 |
| Pines | 76.2 | 64.0 | 88.9 | 80.0 | 88.9 | 80.0 | 46.7 | 42.0 | 52.4 | 44.0 | 48.6 | 36.0 |
| Eucalyptus | 76.2 | 64.0 | 82.6 | 76.0 | 84.8 | 78.0 | 40.7 | 22.0 | 53.3 | 16.0 | 44.1 | 30.0 |
| Water | 96.2 | 100.0 | 100.0 | 100.0 | 100.0 | 98.0 | 93.9 | 92.0 | 98.0 | 96.0 | 96.1 | 98.0 |
| Overall accuracy | 82.0 | | 90.0 | | 91.3 | | 48.3 | | 49.3 | | 48.0 | |
| Kappa coefficient | 0.784 | | 0.88 | | 0.896 | | 0.38 | | 0.392 | | 0.376 | |

S1, Sentinel-1; S2, Sentinel-2; UA, user accuracy; PA, producer accuracy

Fig. 3 The best-classified maps of native and non-native habitats. **a** AVIRIS-NG with SVM, **b** S1 + S2 datasets with RF, **c** Sentinel-2 with SVM, and **d** Sentinel-2 with RF



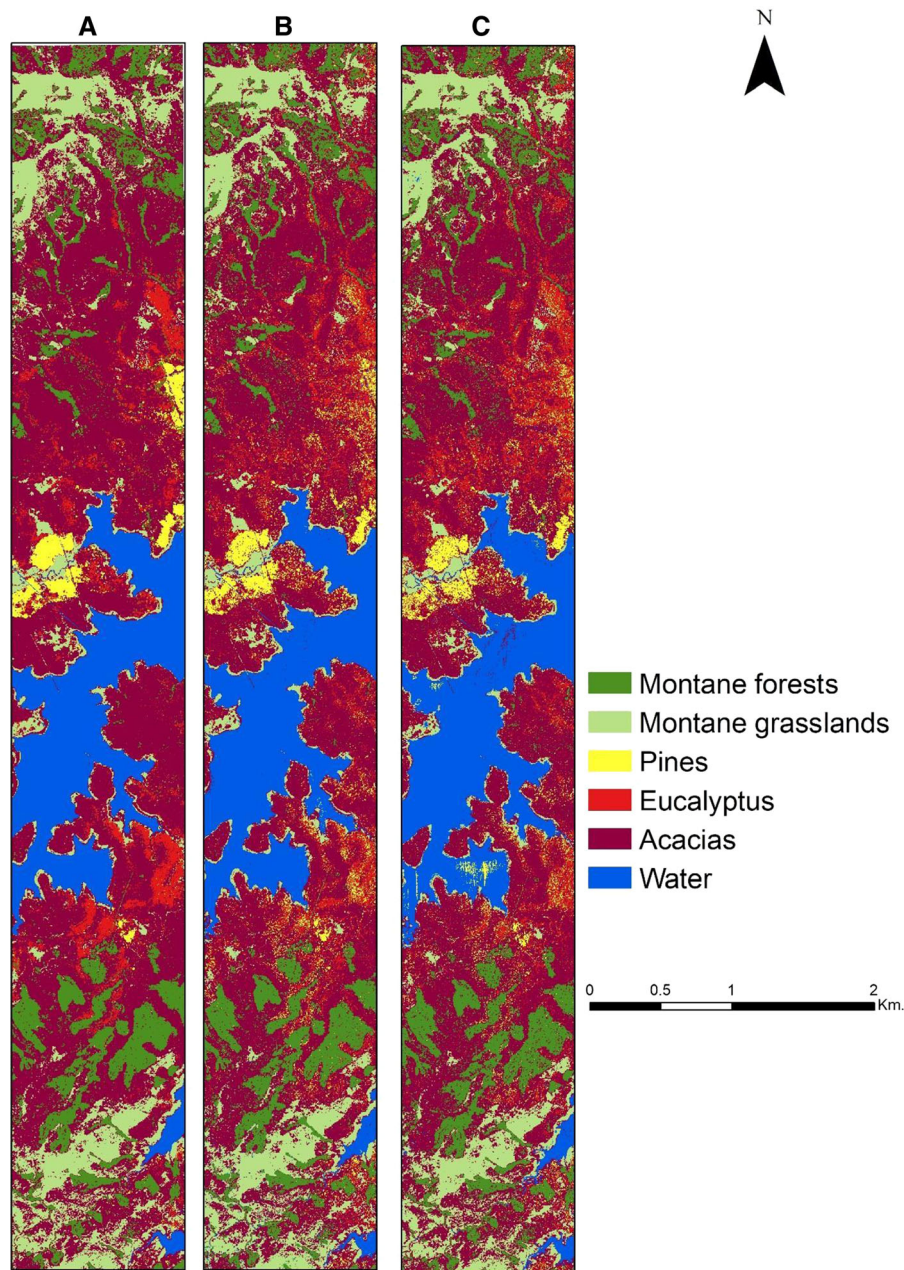


Fig. 4 The AVIRIS-NG classified maps of native and non-native habitats. **a** Classified with SVM, **b** RF, **c** CART

classification (60%) followed by SVM classification (58%) and then CART classification (38%). The S1 + S2 datasets and Sentinel-1 alone produced less than 40% accuracy across all classifiers.

Discussion

The systematic monitoring and mapping of invasive alien species are essential for the conservation and restoration of tropical grasslands. Therefore, spatially-explicit information on native habitats and non-native species is critical for sustainable forest management and forecasting landscape changes into the future.

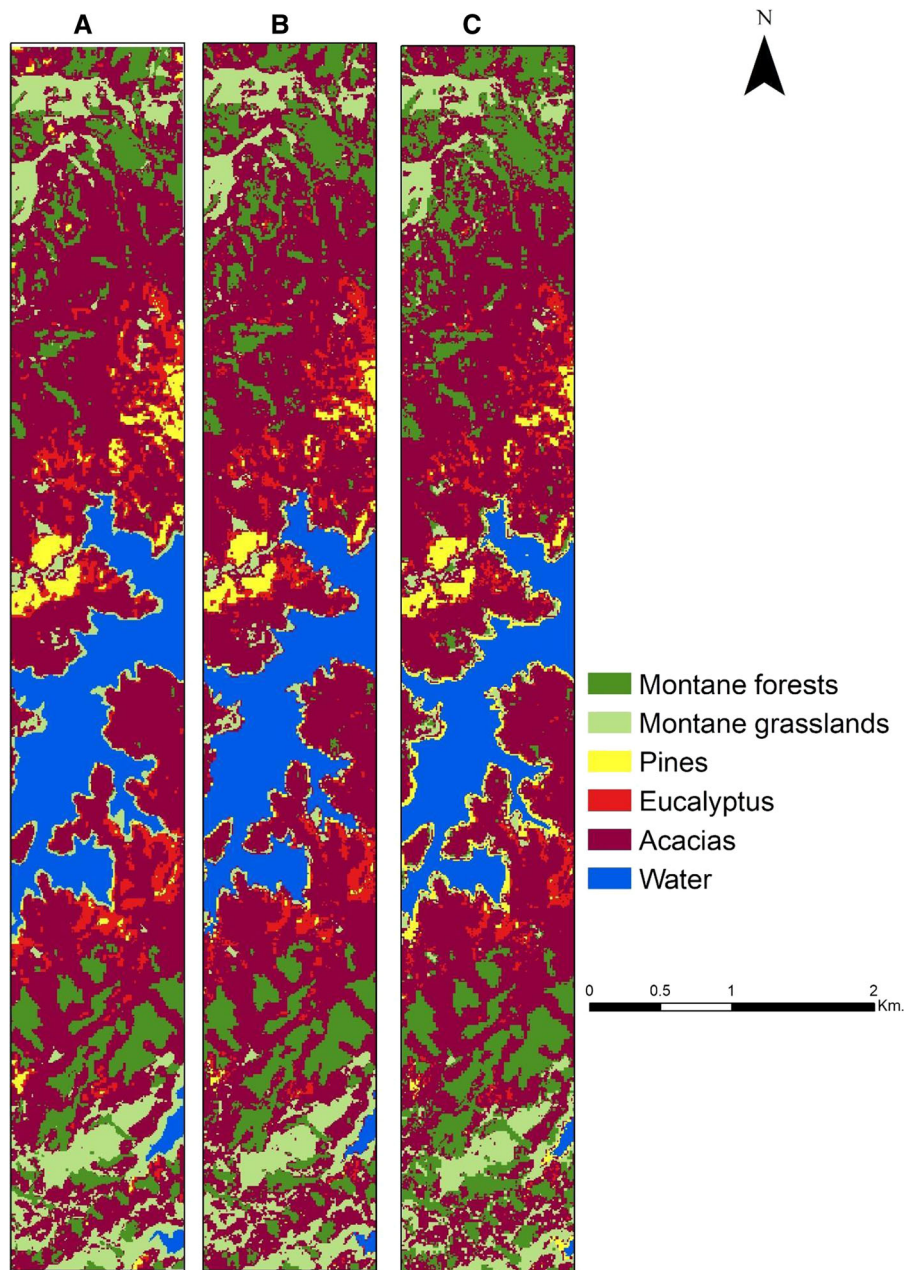


Fig. 5 The Sentinel-2 classified maps of native and non-native habitats. **a** Classified with SVM, **b** RF, and **c** CART

Image data sources

In this study, we found that the hyperspectral dataset (AVIRIS-NG in this study) was the ideal platform for discriminating between native trees and non-native invasive trees with high precision. The AVIRIS-NG dataset accurately delineated the edges of non-native trees and native habitats, likely due to its high spatial

and spectral resolution. The S1 + S2 datasets comprising multispectral and radar data proved to be a reasonable alternative but were only marginally better than using Sentinel-2 images alone. Finally, Sentinel-1 data did not produce sufficient accuracy for classifying and differentiating the invasive species from the native species in the tropical montane habitats. Similar results have been reported in classifying forest types

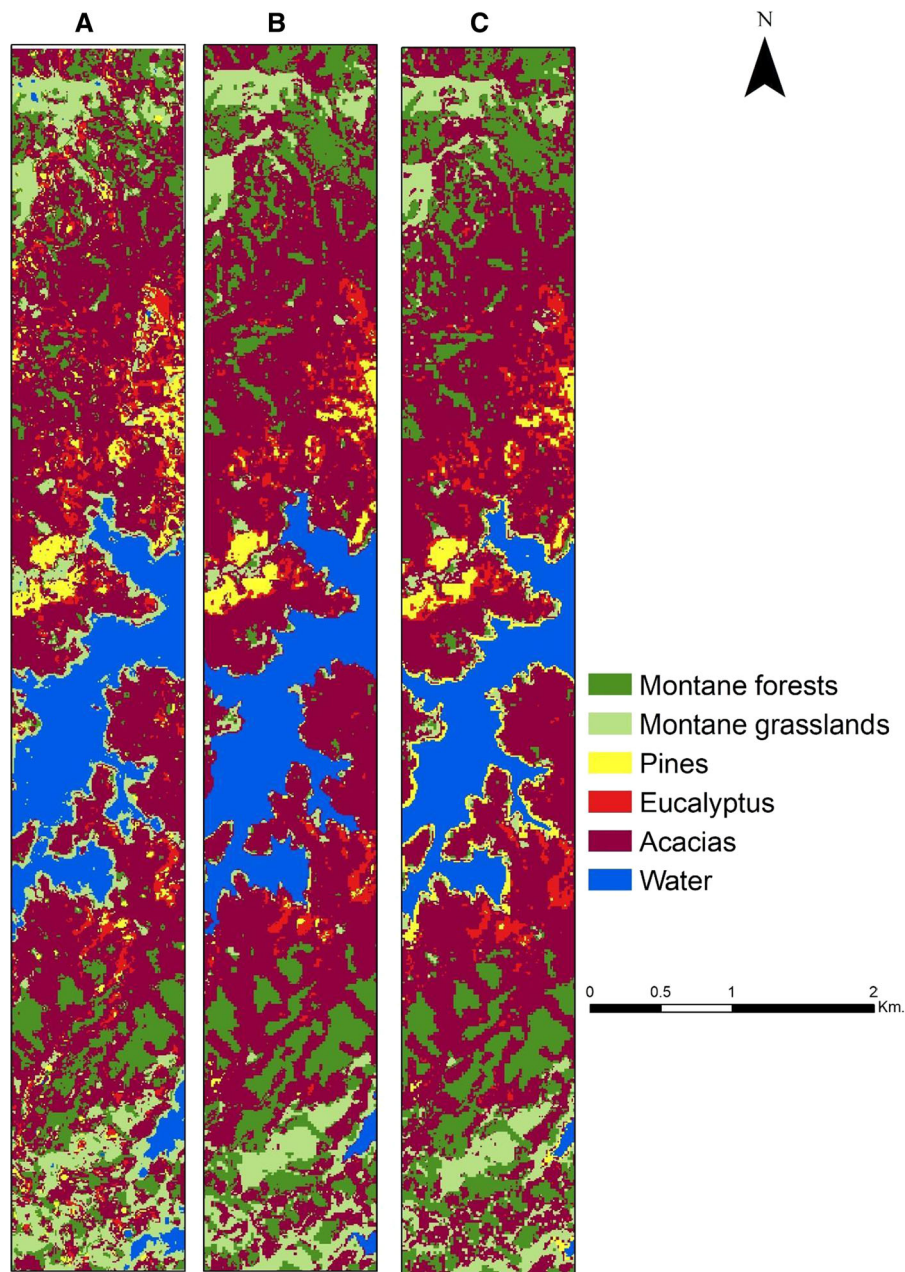


Fig. 6 The Sentinel-1 and Sentinel-2 classified maps of native and non-native habitats. **a** Classified with SVM, **b** RF, and **c** CART

with Sentinel-1 data in these tropical regions. Sentinel-1 data may, however, be probably useful when the Sentinel-2 images are affected by clouds (Erinjery et al. 2018).

Our study supports results obtained by others using hyperspectral remote sensing data for tropical tree species mapping efforts (Hyperion data—30 m spatial resolution 293 and 10 nm spectral resolution; George

et al. 2014; Thenkabail et al. 2004). While Sentinel-1 images have shown to be useful for detecting and classifying water bodies in numerous studies (e.g. Bioresita et al. 2018; Hu et al. 2020), we were unable to map invasive woody species accurately. This may be due to similarities in the backscattering SAR signatures of native and non-native species. Sentinel-1 also has a shorter wavelength (C-band) that may not be

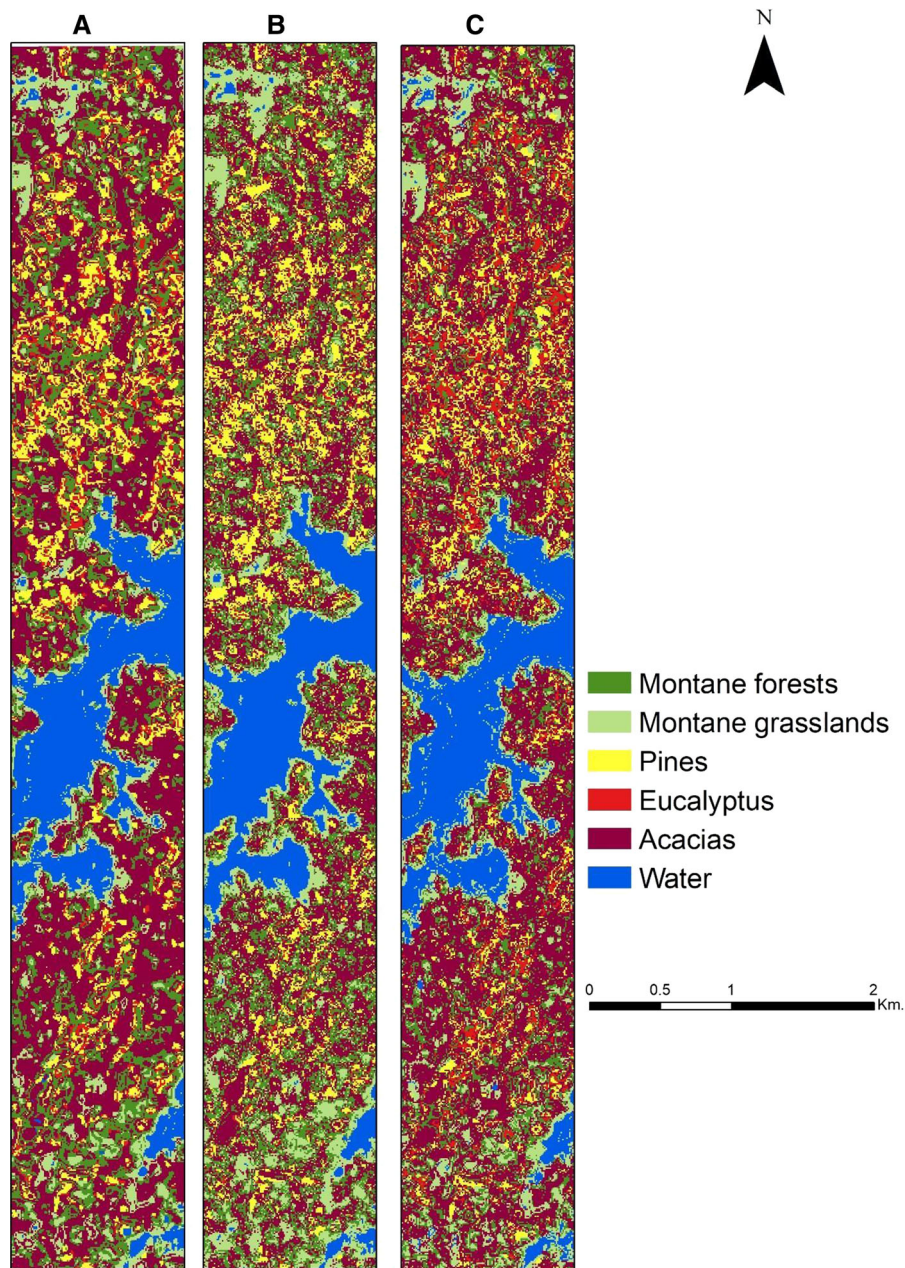


Fig. 7 The Sentinel-1 classified maps of native and non-native habitats. **a** Classified with SVM, **b** RF, and **c** CART

able to differentiate the forest types based on the height and volume information where longer wavelengths (L-band) may be of advantage (Mitchell et al. 2014).

Although our study reports a relatively high accuracy of the combined Sentinel-1 and Sentinel-2 imageries, the difference with the use of only Sentinel-2 was marginal; we believe that this is due

to our specific land cover types since Sentinel-1 image showed similar texture of all three invasive species. This, however, is different from other studies that have reported relatively higher accuracy in the classification from the combined Sentinel-1 and Sentinel-2 imageries for mapping different forest types (Erinjery et al. 2018; Kattenborn et al. 2019), agricultural applications (Veloso et al. 2017), mapping wetlands

Table 2 Accuracy assessment with 25%, 50%, 75% and 100% of training samples

| | OA (25%) | Kappa (25%) | OA (50%) | Kappa (50%) | OA (75%) | Kappa (75%) | OA (100%) | Kappa (100%) |
|------------------------|-------------|----------------|-------------|----------------|-------------|----------------|--------------|-----------------|
| AVIRIS-NG (SVM) | 94.00 | 0.93 | 97.00 | 0.96 | 97.67 | 0.97 | 98.67 | 0.98 |
| AVIRIS-NG (RF) | 81.33 | 0.78 | 85.67 | 0.83 | 86.33 | 0.84 | 88.67 | 0.86 |
| AVIRIS-NG (CART) | 73.00 | 0.68 | 84.00 | 0.81 | 85.33 | 0.82 | 85.67 | 0.83 |
| Sentinel-2 (SVM) | 80.33 | 0.76 | 89.67 | 0.88 | 91.33 | 0.90 | 91.00 | 0.89 |
| Sentinel-2 (RF) | 89.00 | 0.87 | 90.67 | 0.89 | 91.00 | 0.89 | 91.00 | 0.89 |
| Sentinel-2 (CART) | 87.67 | 0.85 | 89.00 | 0.87 | 89.67 | 0.88 | 86.33 | 0.84 |
| Sentinel-1&2 (SVM) | 65.67 | 0.59 | 72.67 | 0.67 | 80.00 | 0.76 | 82.00 | 0.78 |
| Sentinel-1&2 (RF) | 89.33 | 0.87 | 89.00 | 0.87 | 90.67 | 0.89 | 90.00 | 0.88 |
| Sentinel-1&2 (CART) | 87.67 | 0.85 | 90.33 | 0.88 | 88.33 | 0.86 | 91.33 | 0.90 |

OA overall accuracy

(Slagter et al. 2020), extracting urban extents (Iannelli and Gamba 2019), and delineating water bodies (Ahmad et al. 2019).

Classification algorithms

We observed the highest classification accuracy when using the SVM algorithm for classifying the hyperspectral data relative to RF and CART algorithms. This is in agreement with several other studies that have also observed SVM to be highly suitable for classifying hyperspectral data (Burai et al. 2015; George et al. 2014; Lim et al. 2019; Raczko and Zagajewski 2017). Both RF & SVM algorithms produced good classification accuracy with Sentinel-2 images compared to CART, as also observed by Lu et al. (2018). In our tests, we observe that CART is reasonably accurate, but some grassland and acacia invasions around the water bodies tended to be misclassified with pine trees; this was also the issue with a S1 + S2 classification. Although SVM performed well with Sentinel-2 data (91%), it produced low accuracy for the classification with the S1 + S2 datasets (82%). This is perhaps because the SVM algorithm does not do well with noisy data typical to Sentinel-1 data, and when target land cover class may have similar backscattered textures.

While Sentinel-2 did fairly well in detecting invasive species and natural habitats, we found that the combination of AVIRIS-NG data with an SVM

classification model was the only sensor to detect acacia invasion-front with high accuracy. While we are not aware of other studies focusing on detecting invasion fronts, AVIRIS-NG data have been successful for mapping mangroves to the species (Chaube et al. 2019) and for crop type identifications (Salas et al. 2020).

In this research, we utilized medium (S1 and S2—10 m) and high spatial resolution (AVIRIS-NG—4 m) images for IAS mapping. We were, therefore, successful in mapping the species at the pixel scale. We suspect that sub-pixel classification methods may be more suitable for data with spatial resolution images over 30 m.

Influence of the number of ground truth points

We observed that a limited number of ground truth points (~ 25% of our points) were sufficient for obtaining high accuracies when using hyperspectral data (AVIRIS-NG). However, multispectral data (Sentinel-2) seemed to require a higher density of ground truth points (> 50%) to get reasonable accuracies. If the study area is large, Sentinel-2 data might need three times the number of ground truth points than that required to classify a similar hyperspectral image. Conversely, hyperspectral data (AVIRIS-NG) is not available for all places, but data can be captured using a UAV-borne hyperspectral sensor. Using an

UAV, however, might require significant effort and investment for large study regions.

Influence of spatial and spectral resolution

We believe that the higher spatial resolution of AVIRIS-NG dataset over S1 and S2 provides a distinct advantage in distinguishing smaller native forest patches from non-native trees. We found that the S1 and S2 datasets were not able to detect acacia invasion due to the lower spatial resolution compared to AVIRIS-NG dataset.

The fine spectral resolution of AVIRIS-NG (spanning 425 spectral bands) compared to Sentinel-2 (8

spectral bands) helped discriminate between individual native and invasive species. Other studies have shown similarly high species-level classification accuracies in complex landscapes (George et al. 2014; Skowronek et al. 2017). Spectral signature overlaps in the Sentinel-2 dataset resulted in lower average accuracy in detecting pines and eucalyptus similar to findings from Pu et al. (2012) and Marshall and Thenkabail (2015).

Inferences for conservation managers

Mapping and distinguishing native trees from non-native trees is an essential task for land managers for

Table 3 Recommendations for conservation managers

| Data | Best classification algorithms | Possible applications | Advantages | Limitations |
|---|--------------------------------|---|---|--|
| Hyperspectral data (AVIRIS-NG or UAV with Hyperspectral sensor) | SVM | <ol style="list-style-type: none"> 1. Invasive species mapping 2. Forest species mapping 3. Invasion front detection | <ol style="list-style-type: none"> 1. Require limited ground truth points for classification 2. High precision outputs | <ol style="list-style-type: none"> 1. Applicable for smaller areas 2. Presently, airborne and satellite data are not available for all the areas 3. Capturing and processing UAV data might be time-consuming 4. The data needs to be uploaded to the Google Earth Engine (GEE) platform |
| Multispectral data (Sentinel-2) | SVM & RF | <ol style="list-style-type: none"> 1. Invasive species mapping | <ol style="list-style-type: none"> 1. Data are available globally 2. Data are already available on the GEE platform 3. Suitable for larger areas | <ol style="list-style-type: none"> 1. Needs a high-density of ground truth points 2. Limited accuracy compared to hyperspectral data 3. Difficult to classify small patches of forests 4. Invasion front mapping not possible at the species level |
| Fused dataset—radar and multispectral (S1 + S2) | RF | <ol style="list-style-type: none"> 1. Invasive species mapping | <ol style="list-style-type: none"> 1. Data are freely available in GEE globally | <ol style="list-style-type: none"> 1. Needs more ground truth points 2. Limited accuracy 3. Radar data contains backscatter noise or similar backscatter signature for all the species 4. Invasion front mapping not possible at the species-level |
| Radar data (Sentinel-2) | RF | <ol style="list-style-type: none"> 1. Water bodies extraction | <ol style="list-style-type: none"> 1. Data are freely available in GEE across the world 2. Water bodies can be extracted during the monsoon season | <ol style="list-style-type: none"> 1. Invasive species and forest mapping are not possible |

conserving native and endemic species, assessing fire risks, and impacts on ecosystem services. We found that airborne hyperspectral imagery provides the best solution for detecting invasive species and the invasion front. However, the availability of hyperspectral images is a major limiting factor for most locations. These data were only available only for 12 sq. km. of our study area. Further, procuring such data can be prohibitively expensive. Where these data are not available, we recommend using Sentinel-2 satellite data with SVM or RF algorithms as it offers a reasonable compromise between accurately discriminating native and non-native trees while allowing the mapping of large spatial extents. Sentinel-2 images, however, require a high-density of ground truth points and may still be unsuitable for mapping young invasion because of the limited spatial and spectral resolution. Conservation managers may also consider using RapidEye images with an object-oriented classification approach if they need to map a mixed-species invasion front, and do not need to discriminate among the invasive species along that front (Arasumani et al. 2020). If the invasion front requires constant monitoring in a smaller study area, an unmanned aerial vehicle with a hyperspectral sensor may be indispensable. Finally, we also recommend that conservation managers and restoration NGOs use the online Google Earth Engine (GEE) platform because of the high processing power required for processing high spatial and spectral resolution data. We provide the GEE code used for this study in Appendix 1. We have detailed our recommendations for conservation managers in Table 3.

In this research, we demonstrate the advantage of remotely-sensed hyperspectral, multispectral, and radar data for mapping, and distinguishing native and non-native invasive species using AVIRIS-NG, Sentinel-1 and Sentinel-2 datasets. Although the high-resolution, hyperspectral AVIRIS-NG dataset proved superior, we were severely constrained by its spatial availability. The multispectral Sentinel-2 data, on the other hand, were useful in identifying native and non-native trees across a large landscape. Combined with the processing power of the GEE platform, this study demonstrates the opportunity for species-specific discrimination of invaded landscapes that can be replicated across the globe.

Acknowledgements This research was supported by an IISER-Tirupati postdoctoral research fellowship to Arasumani M and a grant from the Ministry of Environment, Forest and Climate Change, Government of India (Grant F. No. 19-22/2018/RE) to V.V. Robin. We thank Tamil Nadu forest department and APCCF & Field Director, Mudumalai Tiger Reserve for logistics and research permits. We thank Vasanth Bosco for extensive discussions and feedback. We thank Ecology and Evolution lab members at IISER-Tirupati for comments and feedback on the manuscript.

References

- Ahmad SK, Hossain F, Eldardiry H et al (2019) A fusion approach for water area classification using visible, near infrared and synthetic aperture radar for South Asian conditions. *IEEE Trans Geosci Remote Sens* 58(2471):2480
- Alempath M, Rice C (2008) *Nilgiritragus hylocrius*. IUCN Red List Threat Species 2008:e.T9917A13026736. <https://doi.org/10.2305/IUCN.UK.2008.RLTS.T9917A13026736.en>
- Allan DG, Harrison JA, Navarro R et al (1997) The impact of commercial afforestation on bird populations in Mpumalanga Province, South Africa—insights from bird-atlas data. *Biol Conserv* 79:173–185
- Arasumani M, Bunyan M, Robin VV (2021) Opportunities and challenges in using remote sensing for invasive tree species management, and in the identification of restoration sites in tropical montane grasslands. *J Environ Manag* 280:111759
- Arasumani M, Khan D, Das A et al (2018) Not seeing the grass for the trees: timber plantations and agriculture shrink tropical montane grassland by two-thirds over four decades in the Palani Hills, a Western Ghats Sky Island. *PLoS ONE* 13:e0190003
- Arasumani M, Khan D, Vishnudas C et al (2019) Invasion compounds an ecosystem-wide loss to afforestation in the tropical grasslands of the Shola Sky Islands. *Biol Conserv* 230:141–150
- Armstrong A, Van Hensbergen H (1995) Effects of afforestation and clearfelling on birds and small mammals at Grootvadersbosch, South Africa. *S Afr For J* 174:17–21
- Awad MM (2018) Forest mapping: a comparison between hyperspectral and multispectral images and technologies. *J For Res* 29:1395–1405
- Bioresita F, Puissant A, Stumpf A et al (2018) A method for automatic and rapid mapping of water surfaces from sentinel-1 imagery. *Remote Sensing* 10:217
- Breiman L (2001) Random forests. *Mach Learn* 45:5–32
- Breiman L, Friedman J, Stone CJ et al (1984) *Classification and regression trees*. CRC Press
- Burai P, Deák B, Valkó O et al (2015) Classification of herbaceous vegetation using airborne hyperspectral imagery. *Remote Sens* 7:2046–2066
- Burges CJ (1998) A tutorial on support vector machines for pattern recognition. *Data Min Knowl Discov* 2:121–167
- Burgess N, Butynski T, Cordeiro N et al (2007) The biological importance of the Eastern Arc Mountains of Tanzania and Kenya. *Biol Conserv* 134:209–231

- Chaube NR, Lele N, Misra A et al (2019) Mangrove species discrimination and health assessment using AVIRIS-NG hyperspectral data. *Curr Sci* 116:1136
- Chen B, Xiao X, Ye H et al (2018) Mapping forest and their spatial-temporal changes from 2007 to 2015 in tropical hainan island by integrating ALOS/ALOS-2 L-Band SAR and landsat optical images. *IEEE J Sel Top Appl Earth Observ Remote Sens* 11:852–867
- Congalton RG (1991) A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sens Environ* 37:35–46
- Congalton RG, Green K (2019) Assessing the accuracy of remotely sensed data: principles and practices. CRC Press
- Daehler CC (2005) Upper-montane plant invasions in the Hawaiian Islands: patterns and opportunities. *Perspect Plant Ecol Evol Syst* 7:203–216
- de Abreu RC, Durigan G (2011) Changes in the plant community of a Brazilian grassland savannah after 22 years of invasion by *Pinus elliottii* Engelm. *Plant Ecol Divers* 4:269–278
- Delalay M, Tiwari V, Ziegler AD et al (2019) Land-use and land-cover classification using Sentinel-2 data and machine-learning algorithms: operational method and its implementation for a mountainous area of Nepal. *J Appl Remote Sens* 13:014530
- Donnelly J (1985) Community structure of epigeaic ants in a pine plantation and in newly burnt fynbos. *J Entomol Soc S Afr* 48:259–265
- Dye P, Jarman C (2004) Water use by black wattle (*Acacia mearnsii*): implications for the link between removal of invading trees and catchment streamflow response: working for water. *S Afr J Sci* 100:40–44
- Erinjery JJ, Singh M, Kent R (2018) Mapping and assessment of vegetation types in the tropical rainforests of the Western Ghats using multispectral Sentinel-2 and SAR Sentinel-1 satellite imagery. *Remote Sens Environ* 216:345–354
- Foody G, Hill R (1996) Classification of tropical forest classes from Landsat TM data. *Int J Remote Sens* 17:2353–2367
- Gaertner M, Novoa A, Fried J et al (2017) Managing invasive species in cities: a decision support framework applied to Cape Town. *Biol Invasions* 19:3707–3723
- George R, Padalia H, Kushwaha S (2014) Forest tree species discrimination in western Himalaya using EO-1 Hyperion. *Int J Appl Earth Obs Geoinf* 28:140–149
- Gorelick N, Hancher M, Dixon M et al (2017) Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens Environ* 202:18–27
- Hu S, Qin J, Ren J et al (2020) Automatic extraction of water inundation areas using Sentinel-1 data for large plain areas. *Remote Sens* 12:243
- Iannelli GC, Gamba P (2019) Urban extent extraction combining sentinel data in the optical and microwave range. *IEEE J Sel Top Appl Earth Obs Remote Sens* 12:2209–2216
- Joshi AA, Sankaran M, Ratnam J (2018) 'Foresting' the grassland: historical management legacies in forest-grassland mosaics in southern India, and lessons for the conservation of tropical grassy biomes. *Biol Conserv* 224:144–152
- Kattenborn T, Lopatin J, Förster M et al (2019) UAV data as alternative to field sampling to map woody invasive species based on combined Sentinel-1 and Sentinel-2 data. *Remote Sens Environ* 227:61–73
- Laurin GV, Liesenberg V, Chen Q et al (2013) Optical and SAR sensor synergies for forest and land cover mapping in a tropical site in West Africa. *Int J Appl Earth Obs Geoinf* 21:7–16
- Lawrence RL, Wright A (2001) Rule-based classification systems using classification and regression tree (CART) analysis. *Photogramm Eng Remote Sens* 67:1137–1142
- Le Maitre DC, Gaertner M, Marchante E et al (2011) Impacts of invasive Australian acacias: implications for management and restoration. *Divers Distrib* 17:1015–1029
- Ledgard N (2001) The spread of lodgepole pine (*Pinus contorta*, Dougl.) in New Zealand. *For Ecol Manag* 141:43–57
- Lele A, Arasumani M, Vishnudas C et al (2020) Elevation and landscape change drive the distribution of a montane, endemic grassland bird. *Ecol Evol* 10:7755–7767
- Lim J, Kim K-M, Jin R (2019) Tree species classification using hyperion and sentinel-2 data with machine learning in South Korea and China. *ISPRS Int J Geo Inf* 8:150
- Lu L, Tao Y, Di L (2018) Object-based plastic-mulched land-cover extraction using integrated Sentinel-1 and Sentinel-2 data. *Remote Sens* 10:1820
- Marshall M, Thenkabail P (2015) Advantage of hyperspectral EO-1 Hyperion over multispectral IKONOS, GeoEye-1, WorldView-2, Landsat ETM+, and MODIS vegetation indices in crop biomass estimation. *ISPRS J Photogramm Remote Sens* 108:205–218
- McConnachie MM, Wilgen BW, Richardson DM et al (2015) Estimating the effect of plantations on pine invasions in protected areas: a case study from South Africa. *J Appl Ecol* 52:110–118
- Melgani F, Bruzzone L (2004) Classification of hyperspectral remote sensing images with support vector machines. *IEEE Trans Geosci Remote Sens* 42:1778–1790
- Mitchell AL, Tapley I, Milne AK et al (2014) C-and L-band SAR interoperability: Filling the gaps in continuous forest cover mapping in Tasmania. *Remote Sens Environ* 155:58–68
- Mooney HA (2005) Invasive alien species: a new synthesis. Island Press
- Mountrakis G, Im J, Ogole C (2011) Support vector machines in remote sensing: a review. *ISPRS J Photogramm Remote Sens* 66:247–259
- Pellikka PK, Lötjönen M, Siljander M et al (2009) Airborne remote sensing of spatiotemporal change (1955–2004) in indigenous and exotic forest cover in the Taita Hills, Kenya. *Int J Appl Earth Obs Geoinf* 11:221–232
- Pu R, Bell S, Meyer C et al (2012) Mapping and assessing seagrass along the western coast of Florida using Landsat TM and EO-1 ALI/Hyperion imagery. *Estuar Coast Shelf Sci* 115:234–245
- Raczko E, Zagajewski B (2017) Comparison of support vector machine, random forest and neural network classifiers for tree species classification on airborne hyperspectral APEX images. *Eur J Remote Sens* 50:144–154
- Richardson D, Van Wilgen B (1986) Effects of thirty-five years of afforestation with *Pinus radiata* on the composition of mesic mountain fynbos near Stellenbosch. *S Afr J Bot* 52:309–315
- Richardson DM (1998) Forestry trees as invasive aliens. *Conserv Biol* 12:18–26

- Rundel PW, Dickie IA, Richardson DM (2014) Tree invasions into treeless areas: mechanisms and ecosystem processes. *Biol Invasions* 16:663–675
- Salas EAL, Subburayalu SK, Slater B et al (2020) Mapping crop types in fragmented arable landscapes using AVIRIS-NG imagery and limited field data. *Int J Image Data Fus* 11:33–56
- Shaharum NSN, Shafri HZM, Ghani WAWAK et al (2020) Oil palm mapping over Peninsular Malaysia using Google Earth Engine and machine learning algorithms. *Remote Sens Appl Soc Environ* 17:100287
- Shimizu K, Ota T, Mizoue N (2019) Detecting forest changes using dense Landsat 8 and Sentinel-1 time series data in tropical seasonal forests. *Remote Sens* 11:1899
- Six LJ, Bakker JD, Bilby RE (2013) Loblolly pine germination and establishment in plantations and grasslands of northern Uruguay. *For Ecol Manag* 302:1–6
- Skowronek S, Ewald M, Isermann M et al (2017) Mapping an invasive bryophyte species using hyperspectral remote sensing data. *Biol Invasions* 19:239–254
- Slagter B, Tsendbazar N-E, Vollrath A et al (2020) Mapping wetland characteristics using temporally dense Sentinel-1 and Sentinel-2 data: a case study in the St. Lucia wetlands, South Africa. *Int J Appl Earth Obs Geoinf* 86:102009
- Thenkabail PS, Enclona EA, Ashton MS et al (2004) Hyperion, IKONOS, ALI, and ETM+ sensors in the study of African rainforests. *Remote Sens Environ* 90:23–43
- Veloso A, Mermoz S, Bouvet A et al (2017) Understanding the temporal behavior of crops using Sentinel-1 and Sentinel-2-like data for agricultural applications. *Remote Sens Environ* 199:415–426
- Wheeler J, Rodriguez-Veiga P, Balzter H, Tansey K, Tate NJ (2017) Forest mapping of the congo basin using synthetic aperture radar (SAR). *Earth Obs Land Emerg Monit* 57. <https://doi.org/10.1002/9781118793787.ch4>
- Zalba SM, Cuevas YA, Boó RM (2008) Invasion of *Pinus halepensis* Mill. following a wildfire in an Argentine grassland nature reserve. *J Environ Manag* 88:539–546
- Zhang W, Brandt M, Wang Q et al (2019) From woody cover to woody canopies: how Sentinel-1 and Sentinel-2 data advance the mapping of woody plants in savannas. *Remote Sens Environ* 234:111465

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.