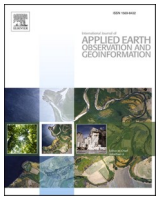
International Journal of Applied Earth Observations and Geoinformation 112 (2022) 102891

Contents lists available at ScienceDirect

International Journal of Applied Earth Observations and Geoinformation

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



A novel method to simulate AVIRIS-NG hyperspectral image from Sentinel-2 image for improved vegetation/wildfire fuel mapping, boreal Alaska

Anushree Badola [a,*], Santosh K. Panda [a,b], Dar A. Roberts [c], Christine F. Waigl [d], Randi R. Jandt [e], Uma S. Bhatt [a]

[a] Geophysical Institute, University of Alaska Fairbanks, Fairbanks, AK 99775, USA
[b] Department of Natural Resources and Environment and Institute of Agriculture, Natural Resources and Extension, University of Alaska Fairbanks, Fairbanks, AK 99775, USA
[c] Department of Geography, University of California, Santa Barbara, CA 93106, USA
[d] International Arctic Research Center, University of Alaska Fairbanks, Fairbanks, AK 99775, USA
[e] Alaska Fire Science Consortium, International Arctic Research Center, University of Alaska Fairbanks, Fairbanks, AK 99775, USA

ARTICLE INFO

Keywords:
Hyperspectral data
Simulation
Spectral library
UPDM
Boreal forest
Endmembers

ABSTRACT

Detailed vegetation maps are one of the primary inputs for forest and wildfire management. Hyperspectral remote sensing is a proven technique for detailed and accurate vegetation mapping. However, the availability of recent hyperspectral imagery in Alaska is limited because of the logistics and high cost involved in its acquisition. In this study, we simulated AVIRIS-NG (Airborne Visible InfraRed Imaging Spectrometer - Next Generation) hyperspectral data from widely available Sentinel-2 multispectral data using the Universal Pattern Decomposition Method (UPDM). The UPDM is a spectral unmixing technique that uses detailed ground spectra of vegetation classes and the Spectral Response Functions of AVIRIS-NG and Sentinel-2 sensors to simulate imagery with the same number of bands and spectral resolution as an AVIRIS-NG image. We simulated three images (each covering an area of 100 km × 100 km) from two ecoregions to test portability of the approach. We collected ground spectra of vegetation and bare ground during summers (2019–2021) using a PSR+ 3500 hand-held spectroradiometer and created a spectral library for this study. The Iterative Endmember Selection (IES) algorithm was used to optimize the spectral library and to select the most representative endmembers for simulation: birch, spruce, and gravel. We validated the simulated hyperspectral imagery by comparing it with available AVIRIS-NG images. The simulated image was visually and spectrally similar to the AVIRIS-NG image (RMSE of 0.03 and 0.02 for birch and spruce spectra, respectively). We applied the Random Forest image classification model to derive detailed vegetation maps from the simulated images. Our vegetation map showed an improvement of 33% in the map accuracy compared to the LANDFIRE EVT map. This study demonstrated an efficient and cost-effective approach to derive detailed vegetation maps at the Sentinel scene scale by simulating hyperspectral images in Google's cloud environment. It offers a novel pathway to generate detailed vegetation and fuel maps for the whole boreal region of Alaska to aid effective forest and fire management.

1. Introduction

Alaska has an area of approximately 78 million hectares covered with boreal forests (Nowacki et al., 2003; U.S. Geological Survey, 2001). These forests extend from the Coast Range in the south to the Brooks Range in the north. Wildland fire is a ubiquitous feature of the boreal forest and in the past two decades (2001–2020) wildfires burned 12.7 million hectares of forest in Alaska (International Arctic Research Center, 2021). Alaska's boreal forest is highly flammable because of the dominance of black spruce, which is highly combustible, and its low-lying canopy structure, which serves as ladder fuel and promotes crown fires and rapid fire spread. In boreal forests, the ground surface is

* Corresponding author.
E-mail addresses: abadola@alaska.edu (A. Badola), skpanda@alaska.edu (S.K. Panda), dar@geog.ucsb.edu (D.A. Roberts), cwaigl@alaska.edu (C.F. Waigl), rjandt@alaska.edu (R.R. Jandt), usbhatt@alaska.edu (U.S. Bhatt).

https://doi.org/10.1016/j.jag.2022.102891
Received 3 March 2022; Received in revised form 27 May 2022; Accepted 23 June 2022
Available online 1 July 2022

covered with feather moss, lichen, and fine fuels that ignite easily in dry conditions (National Park Service, 2021). Fire managers and the research community aim to improve fire management by generating and using improved fire spread models, climate and fuel inputs. Vegetation/fuel maps are one of the key inputs for fire risk assessment and fire spread modeling. Fire managers need accurate vegetation maps to constrain fire spread by locating the potential areas of risk, appropriately allocating suppression resources, and applying fuel treatments.

Remote sensing of vegetation and forest is a proven approach to mapping vegetation type and wildfire fuel distribution (Dudley et al., 2015; Wagner et al., 2018; Xie et al., 2008; Badola et al., 2019; Badola et al., 2021a; Smith et al., 2021). Specifically, multispectral sensors onboard Landsat, Sentinel 2A and 2B, and Terra/Aqua satellites image the entire globe, and their image data are heavily used for vegetation mapping (Dobrinić et al., 2021; Grabska et al., 2019; Mudele and Gamba, 2019). In the USA, Landscape Fire and Resource Management Planning Tools (LANDFIRE) is a shared program between the U.S. Department of Agriculture and the U.S. Department of the Interior (https://landfire.gov/about.php). It provides geospatial products, including vegetation and fuel maps, to state and federal fire management agencies for wildfire mitigation (Reeves et al., 2009). In Alaska, the LANDFIRE Existing Vegetation Type (EVT) map product derived from Landsat image data (9 bands) at 30 m spatial resolution is traditionally used for fire management, fire spread modeling, and risk assessment. The accuracy of the LANDFIRE EVT 2014 map product ranges between 20% and 45% as per accuracy assessment at three sites (Develice, 2012; Smith et al., 2021). The Alaska Center for Conservation Science (ACCS) offers the Alaska Vegetation and Wetland Composite (AVWC) map product, also generated from 30 m Landsat image data, which includes land cover, wetlands, and deep-water maps for Alaska to promote wetlands and deep-water habitat management (Alaska Vegetation and Wetland Composite, 2019).

Imaging spectroscopy or hyperspectral remote sensing provides an opportunity to generate improved vegetation and fuel maps. A hyperspectral remote sensing sensor images the landscapes in hundreds of narrow contiguous bands, making it more effective for vegetation mapping than a multispectral sensor. However, the high dimensionality of the hyperspectral data may reduce the map accuracy due to the Hughes phenomenon (Hsu, 2007). Despite the constraint posed by high data dimensionality, recent studies (Govender et al., 2019; Badola et al., 2021a; Smith et al., 2021) have shown that hyperspectral data provides more accurate and detailed vegetation/species maps than multispectral data for the boreal region of Alaska. The application of hyperspectral data in vegetation mapping is highly effective, but hyperspectral data are not readily available. NASA JPL provides airborne hyperspectral data collected using the state of the art Airborne Visible/Infrared Imaging Spectrometer-Next Generation (AVIRIS-NG) sensor. These data are highly sought for vegetation/tree species mapping (Ahmad et al., 2021; Badola et al., 2021a; Clark et al., 2005; Hati et al., 2020; Salas et al., 2020; Singh et al., 2020; Smith et al., 2021; Zhang, 2014) due to their narrow bandwidth of 5 nm including a wavelength range of 400–2500 nm, meter-scale spatial resolution, and high signal-to-noise ratio. NASA's JPL team recently acquired AVIRIS-NG data over select sites in Alaska as part of the Arctic-Boreal Vulnerability Experiment (ABoVE) airborne campaign, but these acquisitions cover a fraction of Alaska's boreal forest, so the available hyperspectral data is insufficient for any regional scale vegetation/fuel mapping. In order to have hyperspectral image data for the whole boreal forest of Alaska, we conceived the idea of simulating AVIRIS-NG hyperspectral image data from widely available Sentinel-2 image data at Sentinel scene scale (100 km × 100 km) by modifying the approach developed by Badola et al. (2021b) and implementing it in Google's cloud environment for efficient processing.

Simulation of hyperspectral image data is an emerging research area in the field of remote sensing. A few studies have attempted to simulate hyperspectral data from multispectral data and ground spectra of vegetation and soil using a spectral reconstruction approach. Liu et al. (2009) simulated Hyperion data from ALI multispectral data using the Universal Pattern Decomposition Method (UPDM), a sensor-independent spectral unmixing technique (Zhang et al., 2007). The UPDM calculates the proportion of each class in a pixel and uses the spectral response function of the sensors. This approach was further tested by Tiwari et al. (2016) to simulate Hyperion data from ALI data for Land-Use and Land-Cover (LULC) mapping. They successfully simulated 70 Hyperion bands for a test site in Uttarakhand, India. Using UPDM, Badola et al. (2021b) successfully simulated AVIRIS-NG data (332 bands) from Sentinel-2 data for a boreal forest test site near Fairbanks, Alaska. They obtained higher classification accuracy from simulated data (89%) than Sentinel 2 data (78%). In this study, our goal was to simulate the AVIRIS-NG data from the Sentinel-2 image data at Sentinel-2 scene scale for regional scale vegetation mapping and to test the simulation reproducibility across space and time. We had four research objectives:

A. Implement the Iterative Endmember Selection (IES) algorithm to derive the most representative endmember ground spectra for the boreal region of Alaska.

B. Improve upon the Badola et al. (2021b) simulation algorithm to generate AVIRIS-NG hyperspectral image at Sentinel-2 scene scale (100 km X 100 km).

C. Derive detailed vegetation maps from the simulated hyperspectral data using a machine learning classifier and assess model accuracy and portability across space.

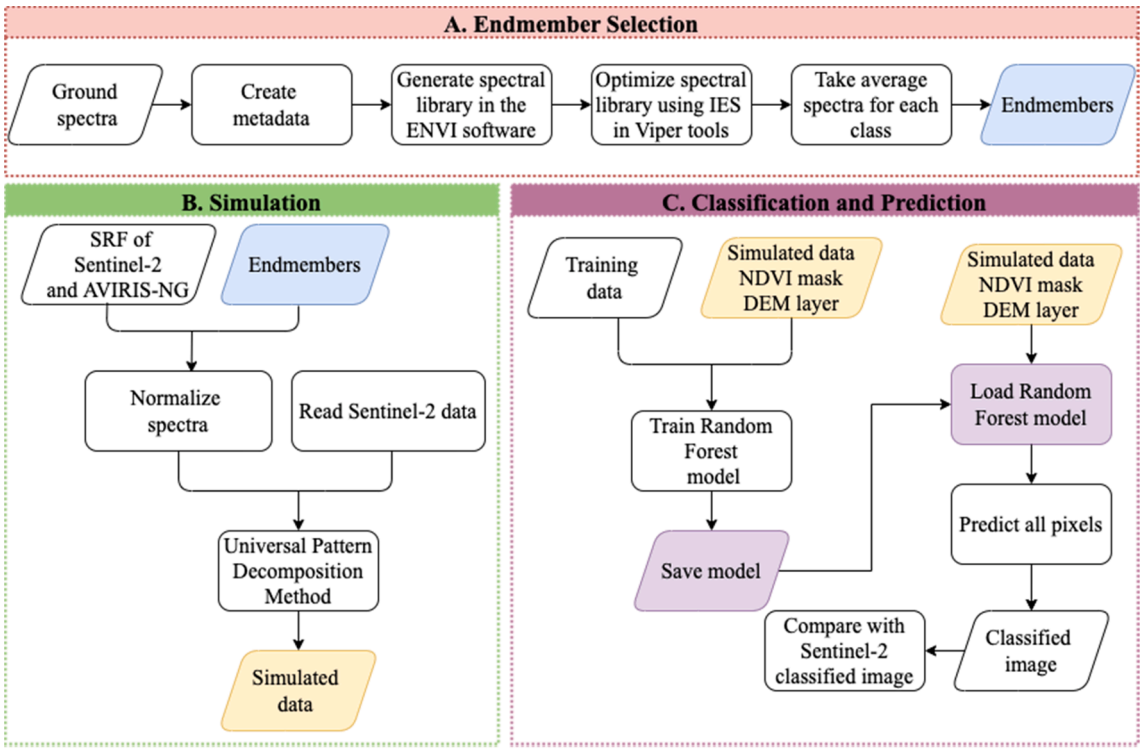
D. Implement the hyperspectral image simulation and vegetation classification algorithms in the Google cloud platform for efficient processing and ease of sharing with the research community.

2. Materials and methods

The methodology is divided into three major phases: A) endmember selection, B) simulation, C) classification and prediction (Fig. 1). We obtained endmembers as the output from the first phase. These were used as input for the simulation phase to simulate the AVIRIS-NG hyperspectral data. We applied a Normalized Differenced Vegetation Index (NDVI) mask and added a Digital Elevation Model (DEM) layer to the simulated data and trained a Random Forest (RF) model using training data collected from the field. Finally, we applied RF on simulated AVIRIS-NG to map vegetation class maps.

2.1. Field data collection

This research required extensive fieldwork to collect ground spectra of vegetation for hyperspectral data simulation as well as vegetation survey for classification of the simulated image. We collected field data (vegetation survey and leaf spectra of all major tree and tall shrub species) over three summers (2019 – 2021). We collected a total of 432 leaf spectra (15 – 20 spectra for each major tree/ tall shrub species) at three sites using a Spectral Evolution® PSR+ 3500 hand-held spectroradiometer (Spectral Evolution Inc., Lawrence, MA, USA). The PSR+ 3500 spectroradiometer provides reflectance data in the range of 350–2500 nm at 1 nm interval, comprising a total of 2151 channels. We collected leaf spectra by holding the optic at about 10 cm distance from the target between 11:00 to 16:00 (local time; local solar noon at 13:56) in sunny, clear-sky weather. We collected branch scale and leaf scale spectra for trees and canopy scale spectra for shrubs (Fig. 2). We also collected NPV spectra using a contact probe for tree bark. We targeted 12 public trails around the Fairbanks city and surveyed vegetation sites using a Garmin handheld GPS device that provides 3 m positional accuracy. At each site, we also recorded the information about canopy cover, vegetation composition, and understory vegetation that helped us in assigning vegetation class to a site. In Fig. 3, the yellow triangles denote the locations of the collected ground data. Table 2 represents all the field data used in this study.



***IES**: Iterative Endmember Selection*
***VIPER**: Visualization and Image Processing for Environmental Research*
***SRF:** Spectral Response Function*
***NDVI**: Normalized Difference Vegetation Index*
***DEM**: Digital Elevation Model*
***AVIRIS-NG**: Airborne Visible InfraRed Imaging Spectrometer - Next Generation.*

Fig. 1. Processing Workflow (A: Endmember Selection; B: Simulation; C: Classification and Prediction).

The AVIRIS-NG scenes were acquired in 2018. Since we used these scenes to assess the spectral quality of simulated hyperspectral data, we used Sentinel scenes closer in date to the AVIRIS-NG scene acquisition dates (Table 1). In the Interior Alaska boreal forest, vegetation change can occur either due to natural succession, insect attack, wildfires, or anthropogenic disturbance such as timber harvesting. During fieldwork, we ensured that there was no evidence of insect outbreak or any major vegetation change at the study sites since 2018.

2.2. Data preprocessing

We used radiometrically and geometrically corrected Sentinel-2A level 1C (Top-of-atmosphere reflectance) data available from European Space Agency (ESA) Copernicus Open Access Hub (European Space Agency, 2014). Table 1 lists the datasets that were used in this study, including the sub-ecoregion where they belong (Nowacki et al., 2003). We used the Sen2cor processor (Louis et al., 2016) available in ESA's Sentinel Application Platform (SNAP) for atmospheric, terrain, and cirrus correction to obtain level 2A surface reflectance data. The size of each scene was 100 km × 100 km. Sentinel-2A data has 13 bands, from which we removed band 1 (coastal aerosol), band 9 (water vapor), and band 10 (SWIR-Cirrus). The visible bands (bands 2, 3, and 4) and NIR band (band 8) have 10 m resolution while the SWIR (bands 11 and 12) and vegetation red edge bands (bands 5, 6, 7, and 8A) have 20 m spatial resolution. We resampled all 20 m bands to 10 m to preserve the best possible spatial resolution and better match the 5 m resolution of AVIRIS-NG. We used atmospherically corrected level 2 AVIRIS-NG data (NASA JPL, 2018) with 425 bands and 5 m spatial resolution to validate the simulated data. We removed bands that contained excessive noise due to atmospheric scattering or are dominated by methane and water vapor absorption. We used the ASTER Global Digital Elevation Model (GDEM) Version 3 (EarthData, 2021) as an additional feature for image classification. It is available at a spatial resolution of 1 arc second (approximately 30 m) (Abrams et al., 2020).

2.3. Building a spectral library of boreal vegetation and endmember selection

We collected a total of 432 spectra (15–20 spectra for each species) of trees/shrubs from the boreal region of interior Alaska. We created a spectral library of all the collected ground spectra using ENVI classic software (Exelis Visual Information Solutions version 5.3, 2010). The PSR+ 3500 hand-held spectroradiometer also records the latitude, longitude, and elevation, along with each target spectrum. We extracted this information from all of the individual spectral files and created a separate metadata file for the spectral library. Furthermore, we did all spectral processing in the Visualization and Image Processing for Environmental Research (VIPER) Tools 2 (beta) software (Roberts et al., 2018).

Library pruning is an important step for creating a spectral library, as it reduces the size of a spectral library and provides the ideal spectra for each endmember. There are different library pruning techniques such as Endmember Average RMSE (EAR), Minimum Average Spectral Angle (MASA), Count-based Endmember Selection (CoB) and Iterative End-

Fig. 2. Vegetation spectra collection in the field; (a) branch scale spectra for birch using a bare fiber optic (b) branch scale spectra for black spruce using a bare fiber optic (c) leaf scale spectra for birch using a bare fiber optic (d) NPV spectra for birch bark using a contact probe.

member Selection (IES). These techniques rely on the square array (Roberts et al., 1997) that stores the information about how an endmember performs when used to unmix other spectra in the same library. The square array is an $n \times n$ grid of pixels where n is the total number of spectra and n along the row denotes the spectrum used for unmixing from other spectra. We can gather information about RMSE, shade fraction, and spectral angle from the square array. For more details on the square array refer to Roberts et al. (1997). In this study we used the IES (Roth et al., 2012; Schaaf et al., 2011) method for library pruning. It calculates the kappa coefficient (McHugh, 2012) to create a subset of spectra which provides the best class separability. IES classifies the entire spectral library using a subset from the original library. Endmembers are iteratively added and removed from the subset until kappa no longer improves. IES has been used for different applications. Roberts, et al. (2015) used IES to discriminate urban surface materials. IES was implemented to map vegetation species (Dudley et al., 2015; Roberts et al., 2015) and for improved burn severity mapping (Fernandez-Manso et al., 2016). The IES algorithm was implemented in the VIPER Tools 2 (beta) software. We ran IES in fully constrained mode (RMSE and fraction constrained) with default parameter settings, i.e., RMSE threshold of 0.025.

After pruning, we got 105 spectra out of 432 for 15 endmember classes (Table 3). We used the average of the spectra as the endmember for simulation. Hence, we ended up with 15 endmembers, including one for gravel. We initiated 15 endmembers as input for the simulation and compared our simulated product with AVIRIS-NG scene by collecting spectra from the known pixel (identified during field work) from both simulated and AVIRIS NG data. We reduced the number of endmembers in each iteration, tried different endmembers combinations (Table S1) for simulation, and simultaneously verified our results until we obtained similar spectra with low RMSE values. For the second iteration, we plotted all 15 endmembers (Figure S1), and removed similar endmember spectra. We ran the simulation model with 9 spectra. In the third iteration, we removed five more endmembers (alder, blueberry, larch, white spruce and asphalt). We kept one spectrum from each deciduous and coniferous class, gravel, and Non-Photosynthetic Vegetation (NPV) spectra (downed trunk). In the fourth iteration, we replaced downed trunk with asphalt and ran the simulation. In the fifth iteration, we used NPV spectra from the ECOSTRESS spectral library (Meerdink et al., 2019). We found a drastic discrepancy between spectra generated from simulated and AVIRIS-NG data in all the iterations. Finally, we used birch, black spruce and gravel for simulation and found similar spectra

Table 1
List of image datasets used in this study.

Data	Scene Identifier	Acquisition Date	Main area covered	Sub-ecoregion
Sentinel-2A	S2AT06WVS	July 22, 2018	Fairbanks	Tanana-Kuskokwim Lowlands
Sentinel-2B	S2BT06WWS	July 24, 2018	East of Fairbanks	Yukon-Tanana Uplands
Sentinel-2B	S2BT06WWU	July 01, 2021	Yukon flats	Yukon-Old Crow Basin
AVIRIS-NG	ang20180723t200207	July 23, 2018	Fairbanks	Tanana-Kuskokwim Lowlands
AVIRIS-NG	ang20170718t202618	July 18, 2017	South East of Fairbanks	Yukon-Tanana Uplands
AVIRIS-NG	ang20190705t192514	July 05, 2019	Yukon flats	Yukon-Old Crow Basin

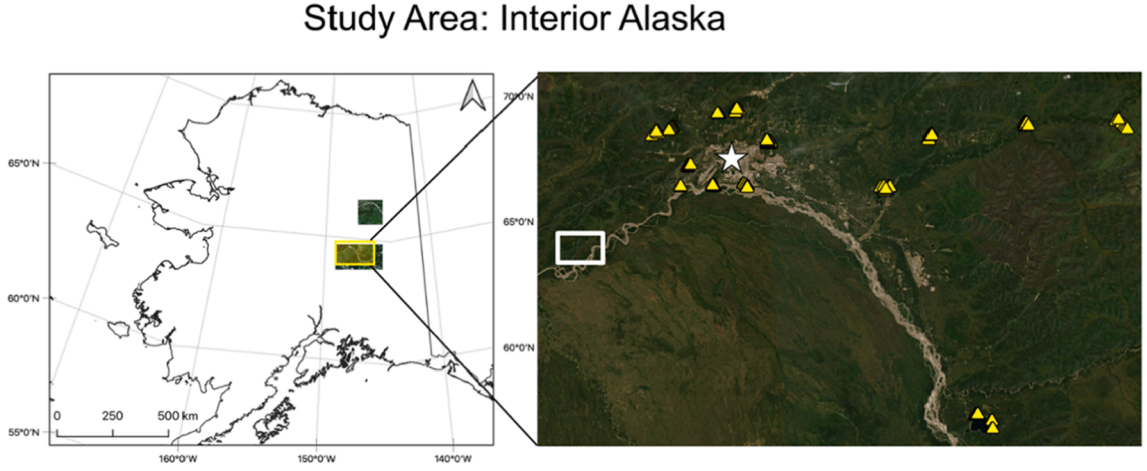


Fig. 3. Study area: Showing the Sentinel-2 scenes used in this study (R: B8, G: B4, B: B3); yellow triangles mark the vegetation survey locations. The star shows the location of Fairbanks, Alaska and the rectangle shows the Bonanza Creek Experimental Forest site (BCEF).

Table 2
List of field data used in this study.

Data	Instrument	Location	Time of Data collection	Data collected
In-situ vegetation survey	Hand held Garmin GPS	12 public trails explored around Fairbanks	Summer, 2021	Vegetation composition, canopy cover, diameter and height
		University of Alaska, Fairbanks campus	Summer, 2019	
		BCEF	Summer, 2020	
Vegetation spectroscopy	PSR+ 3500 Spectroradiometer	University of Alaska, Fairbanks campus	Summer, 2021	Spectra, sample location, vegetation type
	GPS	CPCRW	Summer, 2019 and 2021	
		BCEF	Summer, 2021	

Table 3
List of 15 endmembers obtained through Iterative Endmember Selection library pruning technique.

Endmembers		
Alder	Blueberry	Balsam poplar
Downed trunk	Black spruce	Carex
Dwarf birch	Green grass	Larch
Labrador tea	Birch	Gravel
Asphalt	Wild rose	White spruce

with low RMSE values for spectra generated from simulated and AVIRIS-NG data.

2.4. Simulation of hyperspectral data

The process of hyperspectral data simulation is divided into three steps: ground spectra normalization, weighted fraction coefficients, and hyperspectral data simulation.

2.4.1. Ground spectra normalization

We normalized three endmembers (black spruce, birch and gravel) by convolving them with the Spectral Response Function (SRF) (European Space Agency, 2017) of both Sentinel and AVIRIS-NG sensors. We obtained the SRF for Sentinel-2 data from the Sentinel 2 document library. We calculated the AVIRIS-NG SRF from Full Width at Half Maximum (FWHM) using Gaussian functions (Badola et al., 2021b; Liu et al., 2009).

2.4.2. Weighted fractional coefficient

We used the Universal Pattern Decomposition Method (UPDM) to estimate the proportion of each endmember in every pixel of the image (Badola et al., 2021b; Liu et al., 2009; Tiwari et al., 2016). UPDM is a linear unmixing method that is structured for satellite data analysis (Zhang et al., 2006). It assumes that reflectance at each pixel of an image is a linear mixture of normalized endmembers. The equation expressed in matrix form represents the linear unmixing of three endmembers (b: birch, s: spruce, g: gravel) (Equation (1)).

$$\begin{pmatrix} R_1 \\ R_2 \\ \vdots \\ R_n \end{pmatrix} = \begin{pmatrix} P_{1b} & P_{1s} & P_{1g} \\ P_{2b} & P_{2s} & P_{2g} \\ \vdots & \vdots & \vdots \\ P_{nb} & P_{ns} & P_{ng} \end{pmatrix} \cdot \begin{pmatrix} C_b \\ C_s \\ C_g \end{pmatrix} \quad (1)$$

Where R is the total pixel reflectance, C is the proportion of class, P is the normalized ground reflectance, and n is the band number.

For a multispectral sensor, we can represent Equation (1) as.

$$R_M = P_M C_M \quad (2)$$

C_M is the fraction of coefficients of each endmember in a pixel in the form of a matrix for the whole image. R_M is the matrix with reflectance values from Sentinel 2 multispectral data and P_M is a matrix that contains the reflectance values from the normalized endmembers.

C_M can be calculated from Equation (2) using reflectance from Sentinel 2 data by applying least square method:

$$C_M = (P_M^T.P_M)^{-1}.P_M^T.R_M \quad (3)$$

2.4.3. Hyperspectral data simulation

The spatial resolution of the simulated hyperspectral image will be same as in the Sentinel-2 image, therefore the fraction of coefficients (C_M) will remain the same. We normalized ground spectra (endmembers) using the SRF of the AVIRIS-NG sensor. Hence, we can calculate reflectance values using Equation (1) and Equation (3). The simulation method is discussed in more detail in Badola et al. (2021b). This simulated hyperspectral image has the same number of bands as AVIRIS-NG. Here, in Equation (4), R_H is the reconstructed reflectance values for the simulated hyperspectral image. We write out the simulated hyperspectral image file in GeoTiff file format.

$$R_H = P_H.(P_M^T.P_M)^{-1}.P_M^T.R_M \quad (4)$$

We implemented the hyperspectral image simulation in the Google cloud environment using Python 3 (Python Core Team, 2015) and a Jupyter notebook. We used the following libraries and packages: Pandas to handle the image data in a data frame; Numpy (Harris et al., 2020) to perform the matrix calculations; Rasterio (Gillies et al., 2013) to work with images, especially to read and write the image data. We implemented the algorithm by dividing a Sentinel scene covering an area of 100 km × 100 km into 36 square tiles of 2048 X 2048 pixels.

2.5. Simulated hyperspectral data validation

We validated the simulated data using spectral comparison, statistical analysis and visual interpretation. For spectral and statistical comparison, we extracted pixel spectra for birch and black spruce from AVIRIS-NG image data and the simulated hyperspectral image. These spectra were extracted from the pixels identified in the field. We compared the reflectance values and absorption peaks and visually analyzed the pattern of the spectra. We also calculated the Root Mean Square Error (RMSE) to evaluate the accuracy of the simulated birch and spruce pixel spectra.

We performed visual analysis by generating Colored Infrared (CIR) image using bands with wavelengths 843 nm, 662 nm, and 557 nm as RGB for the AVIRIS-NG and simulated hyperspectral image, and bands with wavelengths 842 nm, 665 nm and 560 nm as RGB for the Sentinel-2 image. We inspected and analyzed different areas of interest based on their visual appearance.

2.6. Image classification

We labeled each survey site to a vegetation class as per Viereck's Alaska Vegetation Classification (Viereck et al., 1992). For vegetation

classes with similar spectra, we merged the classes (e.g., open birch forest and closed birch forest, Figure S2). Using an NDVI threshold of 0.3, we masked out the non-vegetated pixels. We identified 16 vegetation classes, including one 'Other' class that represents non-forested open vegetated areas such as grasslands.

We performed image classification using the 'RandomForestClassifier' function of the scikit-learn () in Python 3 with 500 decision trees and $\sqrt{(425)} \approx 20$ number of features per subset (Breiman, 2001; Pedregosa et al., 2011). We divided the survey data into two sets, a training set and testing set, and trained the classifier on a simulated hyperspectral scene (S2AT06WVS) that had the 16 vegetation classes representative of interior Alaska. The trained model was then used to classify two other simulated hyperspectral scenes to assess the model's portability to other sites.

2.7. Accuracy assessment

We performed the accuracy assessment using F1 scores and Intersection over Union (IoU). Each class has a different number of samples. The imbalance in classes can skew the results in favor of a more abundant class or classes with a greater number of samples and result in biased classification accuracy. In the event of imbalanced classes, classification accuracy is not enough to assess the classifier model performance. We used F1 score and IoU measures to evaluate the model performance. The F1 score is the harmonic mean of the precision and recall of the model, with value ranging from 0 to 1. IoU (also known as the Jaccard index) measures the amount of overlap between the predicted and the actual label. A value of 0 means there is no overlap, while 1 denotes complete overlap. An IoU score greater than 0.5 is considered to be a good prediction.

We also compared our classified output with two available vegetation map products: LANDFIRE EVT (LANDFIRE, 2016) and the Alaska Vegetation and Wetland Composite (AKVWC) (Alaska Vegetation and Wetland Composite, 2019). We also compared the percentage cover of each species with the USDA Tanana Valley State Forest Pilot Inventory (Pattison et al., 2018).

3. Results

In this section, we present the spectral, statistical, and visual comparison of simulated hyperspectral image with AVIRIS-NG and Sentinel 2 image, image classification results, and comparison of our classified vegetation map with two other existing products.

3.1. Spectral and statistical comparison

Simulated spectra accurately captured the key absorption features that were available in AVIRIS-NG data. For spectral and statistical analysis, we removed the 93 bands with noise due to atmospheric scattering and poor radiometric correction, and bands dominated by water vapor and methane absorption (Badola et al., 2021b). Fig. 4 shows the comparison of simulated and AVIRIS-NG spectra for birch and spruce vegetation. In spectra extracted from simulated data, the absorption features were similar to AVIRIS-NG spectra. These absorption features are caused by chlorophyll absorption in the red band (around 690 nm), water absorption in the NIR region, and lignin and cellulose absorption in the SWIR region. Around 1400 nm wavelength, a dip due to water absorption is present in both the spectra generated from simulated and AVIRIS-NG data. In both spectra, the difference in the reflectance values over the infrared region (700 nm – 1400 nm) is relatively small, and their pattern is similar. We achieved an RMSE of 0.03 and 0.02 for birch and spruce, respectively.

Fig. 5 shows the band-to-band correlation between AVIRIS-NG and simulated hyperspectral images. A majority of the bands show high correlation, especially in the NIR and SWIR region. We also calculated the Coefficient of Variation (CoV = standard deviation/ mean) for the

Fig. 4. Spectral Signature of (a) Birch and (b) Spruce.

pixel difference between AVIRIS-NG and simulated hyperspectral images for 332 bands (Fig. 6). The CoV is another metric to quantify and visualize the similarity between the two image data. A lower CoV value suggests higher similarity between the two image data, while a higher CoV value suggests less similarity. In our case, the CoV value ranges from 0.2 to 1.3, with higher CoV values along trails and shaded areas where we expect lower simulation accuracy, as we did not include shade fraction as one of the end members. Both band-to-band correlation and the CoV of pixel differences between AVIRIS-NG and simulated hyperspectral images illustrate the quality of simulated hyperspectral image. The band-to-band correlation is higher than 0.6 for 252 bands, and 98% of the total pixels have a CoV less than 1 suggesting that the simulated image satisfactorily captures the spectral details at pixel scale.

3.2. Visual analysis

The simulated image captures the minute details visible in the AVIRIS-NG image (Fig. 7). AVIRIS-NG data were not available for the bottom left side of the area in Fig. 7 (a) and (b), therefore the region is black. Fig. 7 (a) is an image from Creamer's Field - Migratory Waterfowl Refuge area in Fairbanks. The region shown in a yellow circle is an open field that has trees planted on the boundary. Simulated data captured the trees on the edges quite well. In Fig. 7 (b), the small patches inside the yellow circle are water bodies. These water bodies are accurately captured in the simulated image. In Fig. 7 (c), we overlaid ground surveyed points for birch and spruce on the three images. We can see a clear difference in the image tone and color saturation between the coniferous vegetation (spruce) and deciduous vegetation (birch) in all three images. The area dominated by birch has brighter pixels, while the area with spruce has darker pixels. The simulated image correctly captured this tonal and color differences between birch and spruce forests.

3.3. Image classification

We performed vegetation classification on the simulated data according to the Alaska vegetation classification given by Viereck et al.

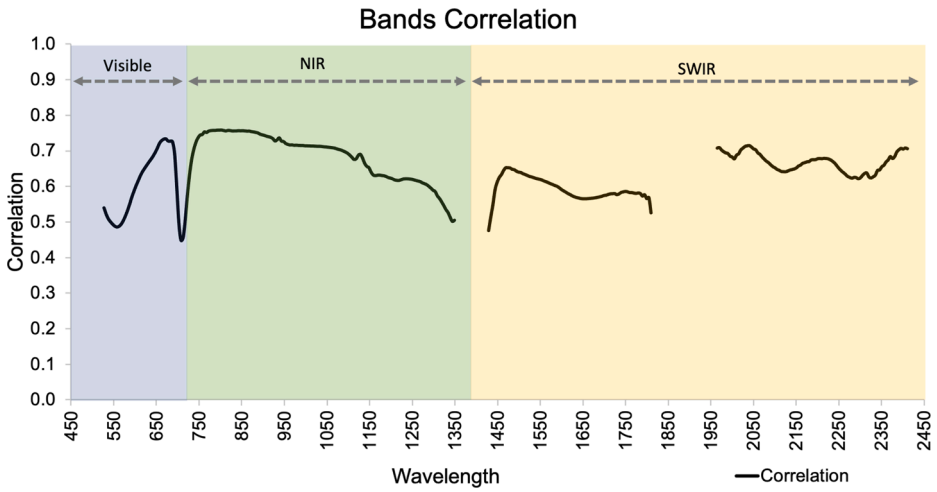


Fig. 5. Band-to-band correlation between AVIRIS-NG and simulated hyperspectral images.

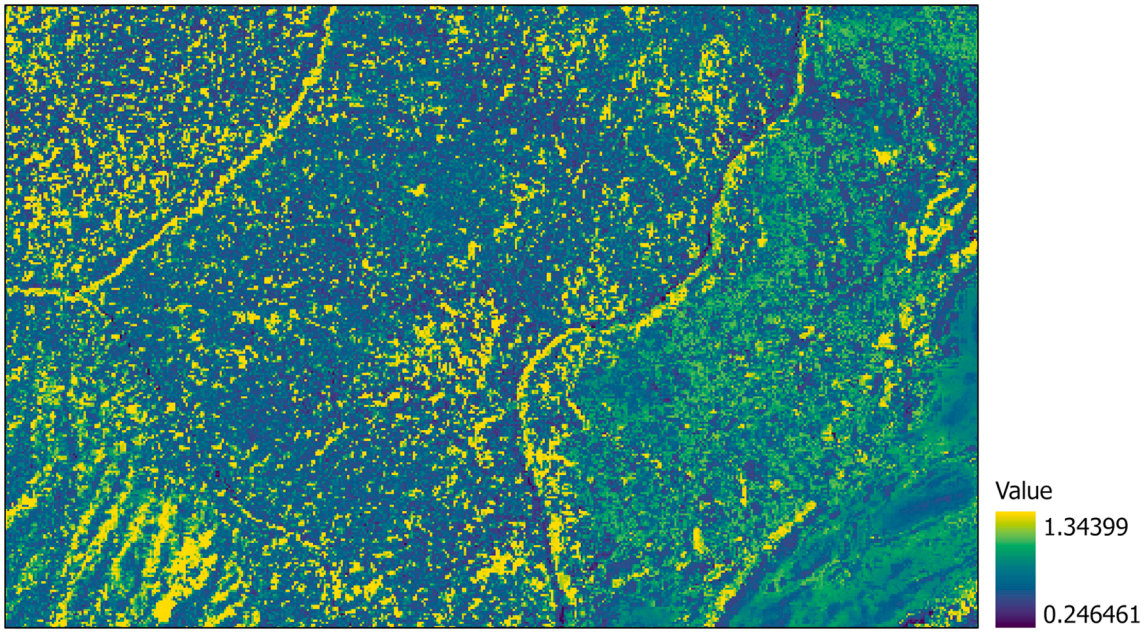


Fig. 6. A map of the Coefficient of Variation (CoV) for the pixel difference between AVIRIS-NG and simulated hyperspectral data.

(1992). Fig. 8 highlights the results of vegetation mapping using the RF classifier. We used NDVI to mask out non-vegetation pixels. There is a separate class for the masked pixels in the legend. Mountains dominate the southern part of the scene; some parts of the mountain and high elevation areas are covered with alpine vegetation. We did not have any training data for alpine vegetation class. Therefore, we included a "Other" class to capture the alpine vegetation and other vegetation (mostly grasses) pixels that are not included in the training. The center of the scene is a lowland area. Most of the pixels are classified into vegetation classes dominated by black spruce.

We had unbalanced test samples; different classes have different numbers of pixels. To overcome class imbalance, we calculated the F1 score and IoU shown in Fig. 9 and Fig. 10 respectively. A high F1 score means the class has performed well. "Open black spruce forest" and "Open balsam poplar forest" had the lowest F1 score, while most classes performed well. Additionally, if the IoU value is 0.5, the class has performed well. In our case, seven classes had an IoU value greater than 0.5, 3 classes had an IoU value very close to 0.5, and the remaining classes had a value between 0.3 and 0.4.

For a test site in BCEF, we generated a vegetation map from Sentinel-2 data and assessed its accuracy using field validation plots. We did not use these plots for training. Out of 31 plots (Fig. 11), 13 plots were correctly mapped with 42% accuracy in the case of Sentinel-2 data, while 20 plots were correctly mapped in the case of simulated classified output, resulting in an accuracy of 65%.

3.4. Process validation

To test the simulation method across space, i.e. for different geographic subregions, we simulated two more scenes (S2BT06WWS, S2BT06WWU) using the same ground spectra and classifier model (trained in scene 1) to predict pixel vegetation class. Scene 2 (S2BT06WWS) covers the area east of scene 1, and scene 3 (S2BT06WWU) covers a part of the Yukon flats region that lies to the north of Fairbanks. Locations of all three scenes are shown in Fig. 3. Scene 2 (S2BT06WWS) falls under the Yukon-Tanana Uplands sub-ecoregion, and scene 3 (S2BT06WWU) covers the Yukon-Old Crow Basin sub-ecoregion. We had a smaller set of ground data for scene 2 that

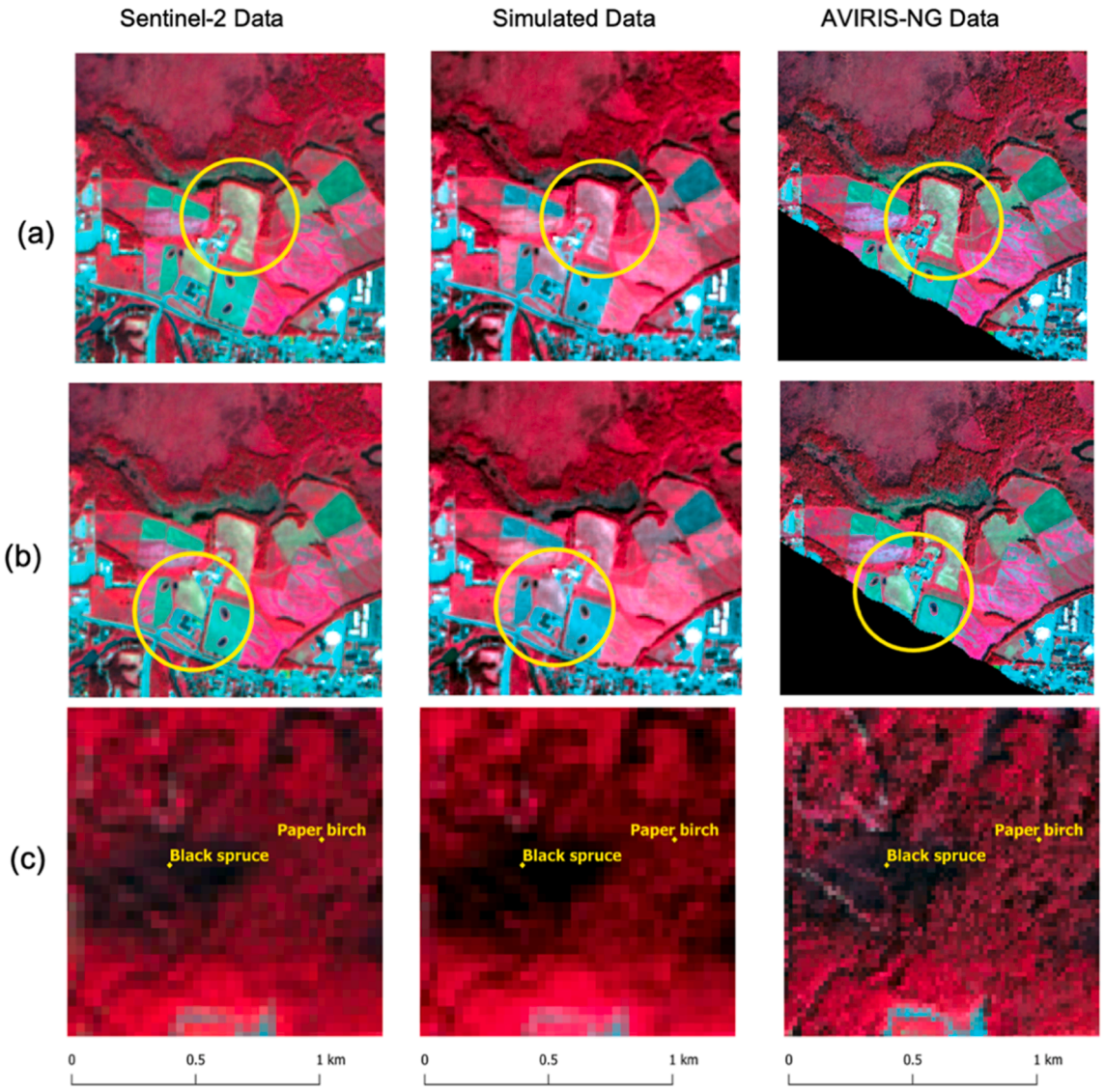


Fig. 7. Visual comparison of the simulated Color Infra-Red image with AVIRIS-NG and Sentinel 2 for different areas: (a) open field (b) water bodies (c) Black spruce and paper birch.

we used for validating our classified vegetation map output. We did not have any ground data for scene 3, so we only performed visual, statistical, and spectral comparison of the simulated image generated from scene 3.

We compared spectra collected from simulated image data and AVIRIS-NG image for scene 2 and scene 3. Figure S3 shows the comparison of the spectra collected from the same pixels in both images. We picked two different sites that represent deciduous and coniferous species. The spectra showed a similar pattern and the same absorption features. The RMSE for deciduous vegetation (Figure S3 (a)) was 0.02, and 0.01 for coniferous vegetation (Figure S3 (b)). We visually compared the simulated data and AVIRIS-NG data. Simulated data has captured the trails and built-up area similar to AVIRIS-NG data (see the yellow circle, Figure S4 (a)). Coniferous and deciduous vegetation are easily distinguishable in simulated data (see the yellow circle, Figure S4 (b)).

Here we have discussed the classified vegetation map accuracy at select points for an area around Twin Bear Chena River to Ridge and Compeau trail east of Fairbanks. We overlaid the ground points over the classified map and visually assessed the classified map product accuracy (Fig. 12). In Fig. 12, points 1, 2, and 5 (from spruce and birch vegetation) are over the blue pixels; the blue pixels in the classified map represent open spruce-paper birch class. Point 4 has paper birch in the ground; the classifier mapped it as the open paper birch forest. Point 6 has white spruce, and the pixels around this point are mapped as the closed white spruce forest. Point 3 has black spruce and moss, and the classified map identified the pixels at and around this point as open black spruce forest and black spruce woodland. The dark gray pixels are wetlands and include all the vegetation growing on marshy areas with seasonal standing water. The gray pixels are mostly around the water bodies (shown as black pixels). Wetlands and vegetation classes present at all six points are correctly mapped in the classified vegetation map.

Using our algorithms, one can simulate hyperspectral data and classify it to generate a vegetation map in 10 h (system configuration: Intel(R) Xeon(R) Gold 5222 CPU @ 3.80 GHz, 3801 MHz, 4 cores with 192 GB RAM). Each simulated hyperspectral image requires a minimum of 190 GB disk space (scene size: 100 X 100 km; pixel size: 10 m; 425 bands). To overcome the dependency on powerful computer systems with large storage, we implemented the entire processing in the Google Cloud Platform (GCP system configuration: 16 CPUs and 60 GB RAM).

4. Discussion

We developed a novel approach to efficiently simulate an AVIRIS-NG

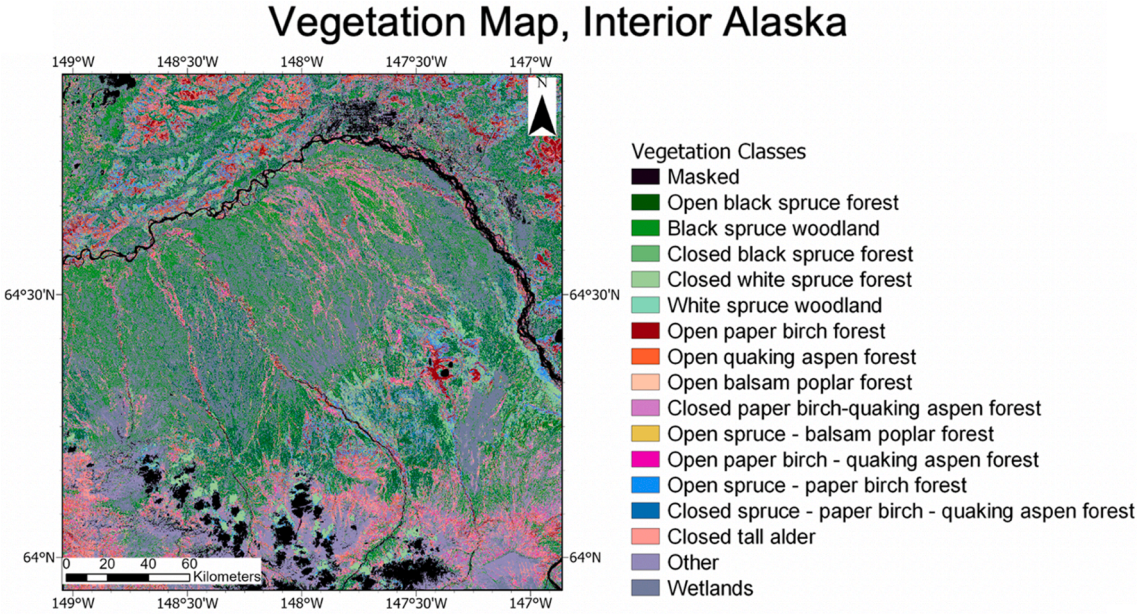


Fig. 8. Vegetation map generated from simulated hyperspectral image (covering an area of 100 km × 100 km) using Random Forest classifier. The masked pixels shown in black include urban areas, water bodies, clouds and cloud shadows.

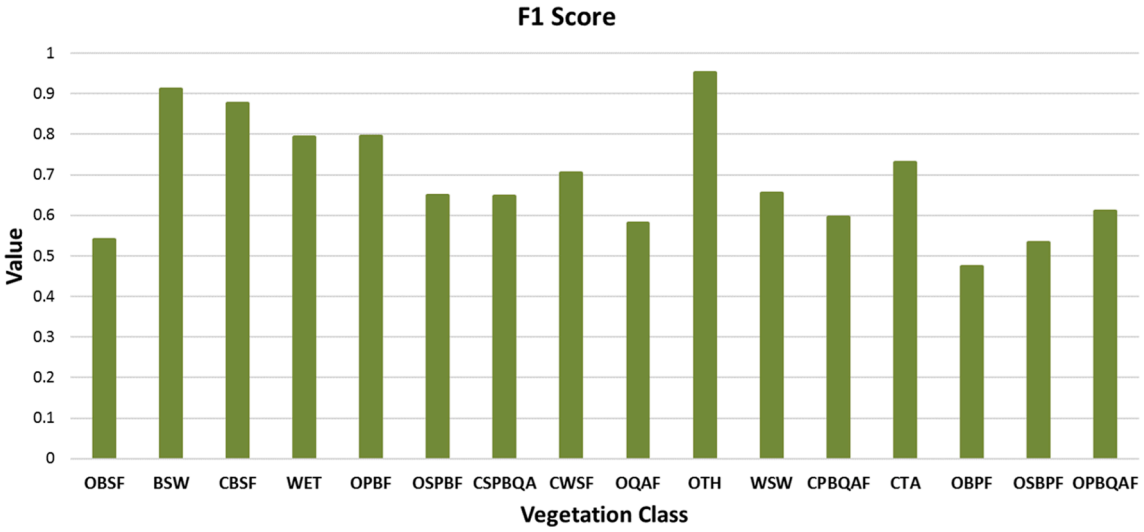


Fig. 9. F1-score for each class.
Open black spruce forest: OBSF; Black spruce woodland: BSW; Closed black spruce forest: CBSF; Wetlands: WET; Open paper birch forest: OPBF; Open spruce - paper birch forest: OSPBF; Closed spruce - paper birch- quaking aspen forest: CSPBQA; Closed white spruce forest: CWSF; Open quaking aspen forest: OQAF; Other: OTH; White spruce woodland: WSW; Closed paper birch-quaking aspen forest: CPBQAF; Closed tall alder: CTA; Open balsam poplar forest: OBPF; Open spruce - balsam poplar forest: OSBPF; Open paper birch - quaking aspen forest: OPBQAF.

hyperspectral image cube from a Sentinel-2 image and subsequently derive a detailed vegetation map covering an area of 100 km × 100 km for the boreal region of Alaska. The simulated images accurately capture the minute landscape features present in the original AVIRIS-NG image. The spectral profile from the simulated image matches the original AVIRIS-NG image in the pattern but differs in reflectance magnitude in the SWIR region. Zhang et al. (2006) and Badola et al. (2021b) also reported a similar difference in reflectance over the SWIR region. Several factors contribute to the spectral differences: a) the difference in the spatial resolution of the two datasets, i.e. Sentinel-2 bands have a spatial resolution of 10 m and 20 m compared to 5 m for AVIRIS-NG, b) the difference in the sensor's altitude and sun and sensor geometry, and c) difference in date and time of acquisition. Also, the missing endmember, Non-Photosynthetic Vegetation (NPV), is contributing to the reflectance difference. NPV has higher SWIR reflectance (Asner, 1998), and depending on the NPV, a vegetation pixel can have lower reflectance in the visible region than in the SWIR region. However, when we added NPV to the model, it reduced model performance significantly. Hence, one of the major limitations of this model is that it cannot handle NPV as one of the endmembers without generating significant errors. However, by excluding NPV, we ended up with lower reflectance in the SWIR region. In reality, a pixel can have a significant NPV fraction, but this model can only handle distinct endmembers. Also, this model is built for the peak of the growing season and might not do well for images acquired during shoulder season; for example, when leaves start to senescence at the beginning of the fall season. Despite the above limitation, our simulation model can accurately capture the key absorption features and spectral details essential for detailed vegetation and

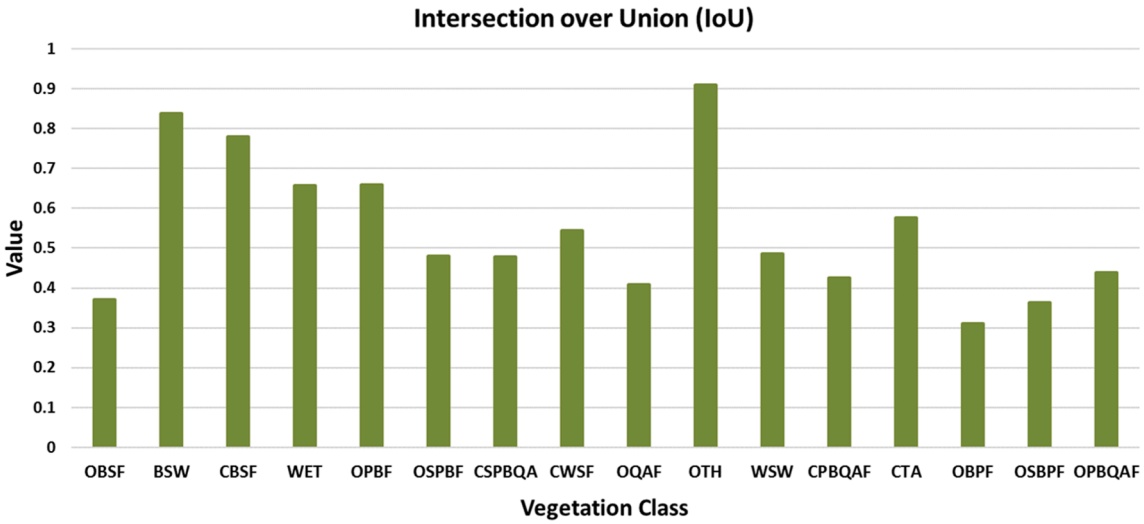


Fig. 10. IoU score for each class.

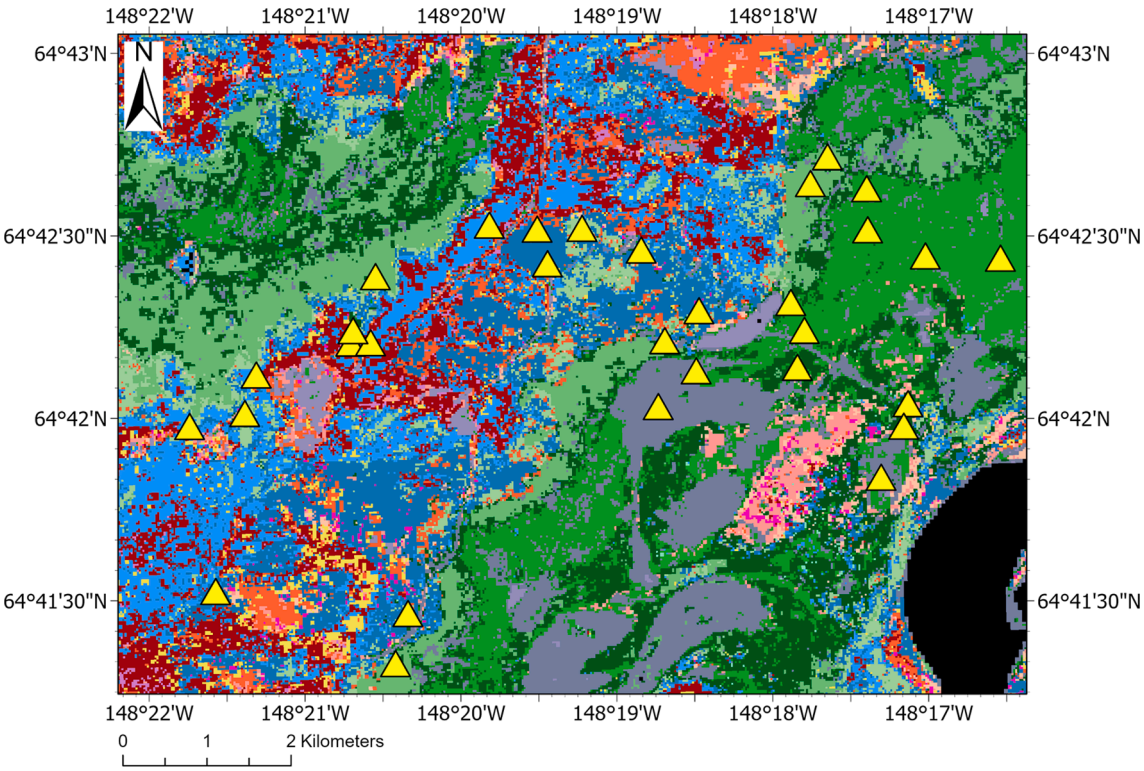


Fig. 11. A vegetation map derived from the Sentinel-2 image for a test site within the Bonanza Creek Experimental Forest (BCEF); yellow triangles mark the field surveyed plot locations used as validation.

landcover mapping, paving the path for on-demand hyperspectral data availability for boreal Alaska.

4.1. Comparison with other map products

We used 31 field surveyed plots in the Bonanza Creek Experimental Forest site, a Long-Term Ecological Research (BCEF LTER) site in Interior Alaska, to assess the accuracy of LANDFIRE EVT and our vegetation map (Fig. 13). Out of 31 plots, 10 plots were mapped correctly in the LANDFIRE EVT 2016 product, giving a product accuracy of 32%, whereas 20 plots were mapped correctly in our vegetation map product, giving a product accuracy of 65%. Table S2 shows the vegetation class present on the 11 plots that were misclassified. The majority of these misclassified plots have similar vegetation. Two plots of closed spruce-paper birch-quaking aspen were incorrectly mapped as open spruce-paper birch, while two other plots were mapped as open quaking aspen. One wetland plot was incorrectly mapped as Black spruce woodland, possibly due to the presence of isolated black spruce and understory vegetation found in black spruce woodland. We believe by merging similar classes and retraining the classifier, the map accuracy can be further improved.

Fig. 14 demonstrates the comparison between our vegetation map and the available LANDFIRE EVT and Alaska Vegetation and Wetland Composite (AKVWC) maps. The rectangle, square, and circle in Fig. 14

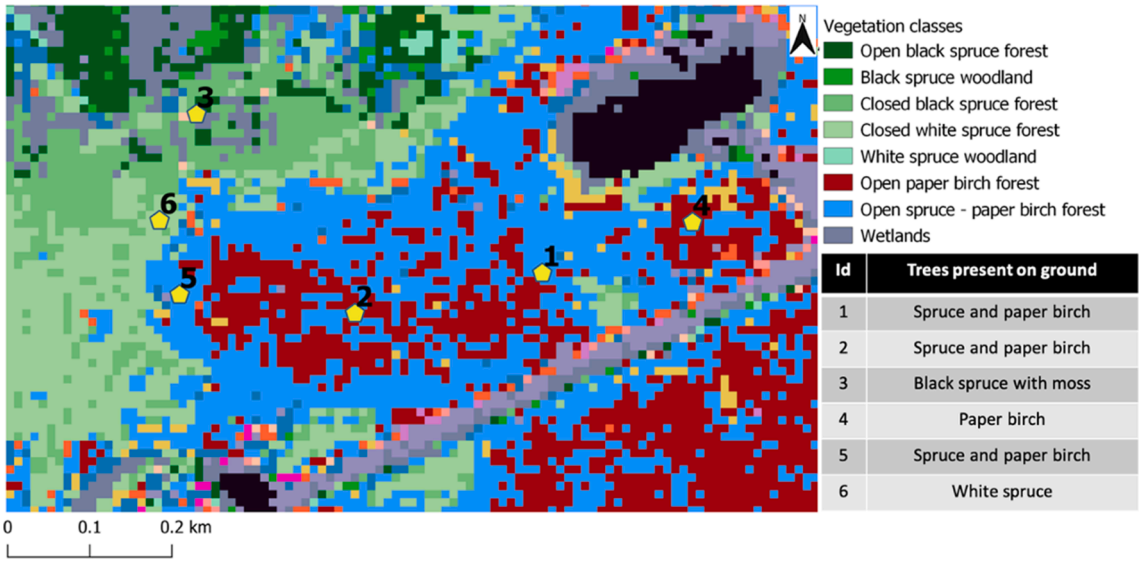


Fig. 12. A vegetation map derived from simulated hyperspectral image for an area around Twin Bear Chena River ridge and Compeau trail; ground data points are presented as yellow pentagon. Tree cover present at the ground point is listed in the table (bottom right).

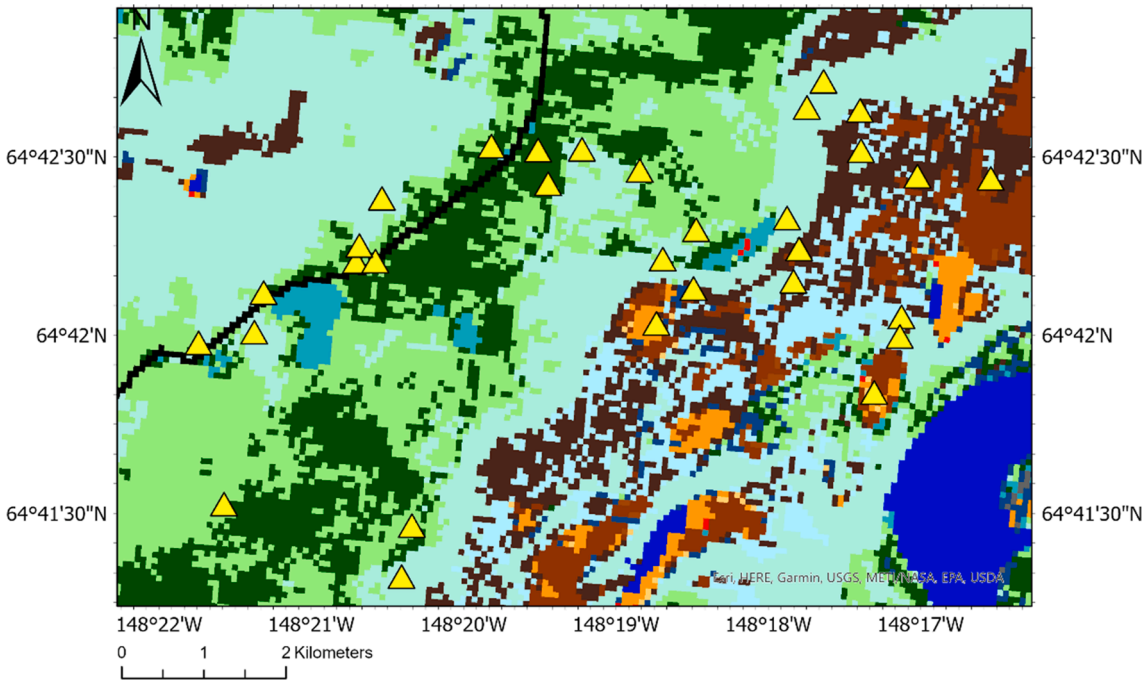


Fig. 13. LANDFIRE EVT Product for the Bonanza Creek Experimental Forest (BCEF) site; yellow triangles mark the field surveyed plot locations.

highlights the locations where we clearly see that the map generated using simulated data is more detailed and accurate in comparison to the other two maps. The region highlighted under the rectangle is dominated by single birch and aspen class in both AKVWC and LANDFIRE EVT maps, while four different birch and aspen vegetation classes are mapped in the simulated classified map: open paper birch forest, open quaking aspen forest, open and closed paper birch, and quaking aspen forest. Similarly, the region inside the circle is dominated by black spruce and white spruce classes in all three cases, but in the case of a simulated classified map, the classes are more detailed. The pixels inside the square are mapped as wetlands in the case of AKVWC and simulated classified maps, while in the LANDFIRE EVT map, most of these pixels are mapped as black spruce-tamarack fen.

The U.S. Department of Agriculture completed a Pilot Inventory for Tanana Valley State Forest, which is covered by the Sentinel scene used in this study. The survey for the Pilot Inventory began in 2014 and was fully implemented between 2016 and 2018. They surveyed 800 plots (around a 7.3 m radius) to get forest type acreage, extrapolating the observations to the entire Tanana Valley State Forest to get total acreage for forest types (Pattison et al., 2018; U.S. Forest Service, 2016). Fig. 15 shows the comparison of forest type acreage estimated from the simulated classified map with estimates from the Pilot Inventory program for five major tree species. For all five tree species, the acreage estimates from our vegetation map are comparable with the estimates from the Pilot Inventory program, and they follow the same pattern as well: i.e. black spruce is the most dominant tree species, followed by birch, with aspen and poplar being the least dominant tree species. Therefore, the forest type acreage we obtained from our map is comparable to the

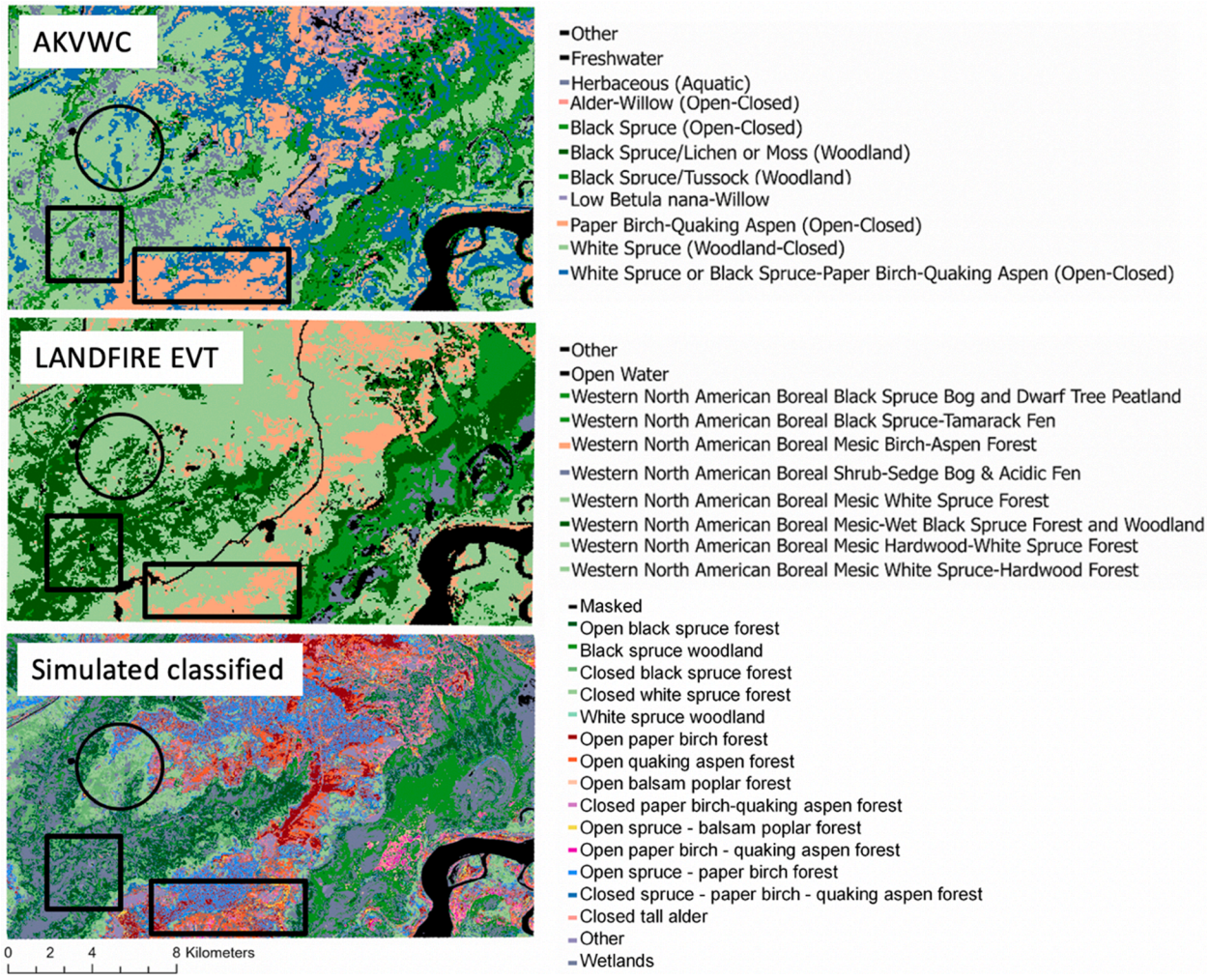


Fig. 14. Comparison of simulated classified map with Alaska Vegetation and Wetland Composite (AKVWC) and LANDFIRE EVT Product.

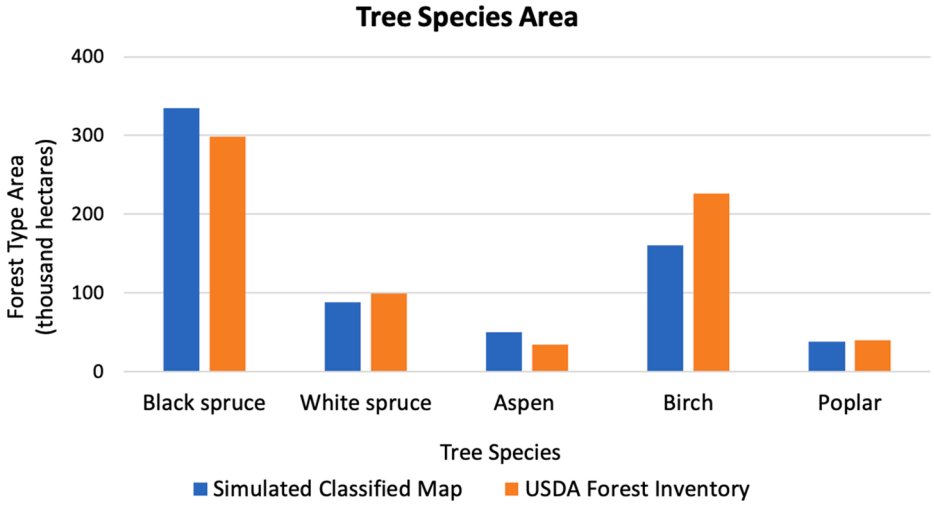


Fig. 15. Comparison of area covered by different forest type reported by the USDA Tanana valley Pilot Inventory program and estimated from our classified map product.

USDA pilot inventory forest acreage.

We used a smaller set of ground data over scene 2 (covering the area east of scene 1) to assess the quality and accuracy of our vegetation map. These points were not used for training the classifier. Fig. 16 compares the simulated classified map with LANDFIRE EVT and Alaska Vegetation and Wetland Composite (AKVWC) maps. In the LANDFIRE EVT map, the black spruce point is sitting on a pixel mapped as Western North American Boreal Mesic White Spruce-Hardwood Forest (bright green

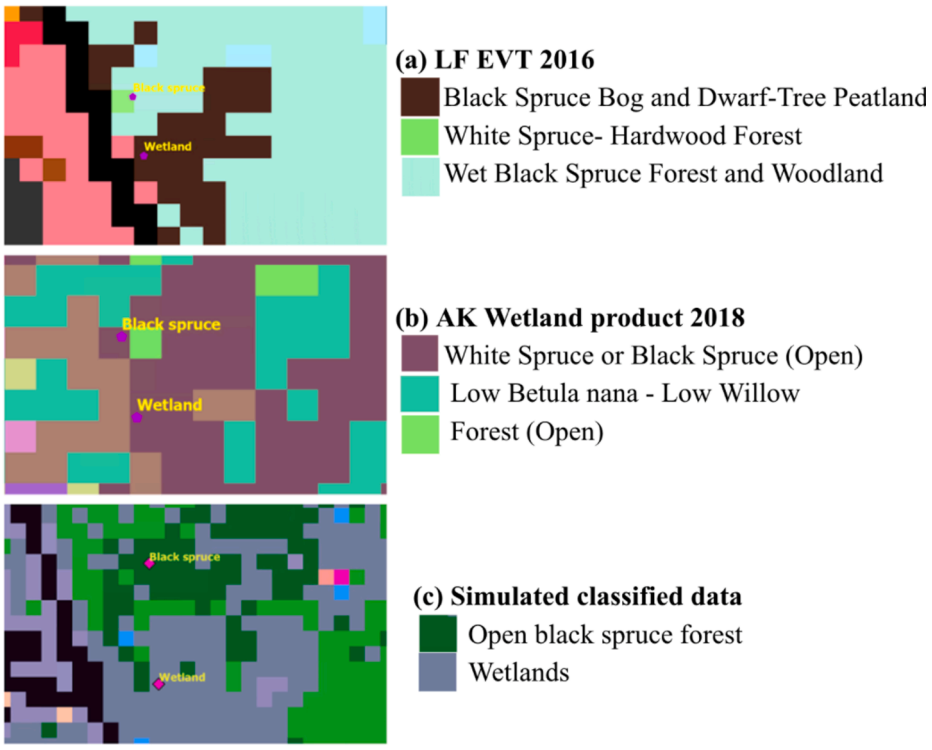


Fig. 16. Comparison of classified map products; ground data points are presented in magenta color. Tree cover present at the ground point is shown on the image in yellow text (a) LANDFIRE EVT Product (b) Alaska Vegetation and Wetland Composite (AKVWC) (c) Simulated classified output.

pixel); in the AKVWC product, the black spruce point is sitting on White Spruce or Black Spruce (Open) (purple pixels). In the simulated classified map, the black spruce point is correctly mapped as open black spruce forest (dark green pixel). For the wetland point, the LANDFIRE EVT product mapped the pixel as Western North American Boreal Black Spruce Bog and Dwarf-Tree Peatland (brown pixel) and AKVWC product mapped the pixel in the same way it mapped for black spruce, i.e. White Spruce or Black Spruce (Open) (purple pixel), while the simulated classified product mapped the area as wetlands (grey pixels). These observations suggest that the simulated classified map correctly identified the open black spruce forest and wetlands pixels with better accuracy than the two existing maps.

We simulated a third scene (scene 3: S2BT06WWU) from a different ecoregion that covers a part of the Yukon Flats region. Figure S5 shows the spectral comparison between simulated data and AVIRIS-NG data. The simulated data accurately captured the absorption features and has a similar reflectance as AVIRIS-NG data: RMSE = 0.02 for deciduous vegetation (Figure S5 (a)) and RMSE = 0.01 for coniferous species (Figure S5 (b)).

The hyperspectral simulation of three different Sentinel 2 scenes and their evaluation demonstrated that we developed an efficient approach for obtaining on-demand hyperspectral data for vegetation/fuel mapping. Our vegetation maps are more detailed and their accuracies are on par with or better than the existing vegetation maps. To effectively implement this approach, it is important to collect field spectra on a clear, sunny day. The field spectra collected on hazy days or cloudy days can degrade hyperspectral simulation results. We also learned that to process the entire Sentinel scene, one needs an expensive system with higher processing power and hardware, so a cloud computing platform like Google Cloud Platform (GCP) (Google, 2022) is a better alternative for efficient and cost-effective processing. GCP reduces the dependency on an expensive local system and the installation of libraries. The only requirements are a browser, a good internet connection, and a Google account to run the codes. More importantly, GCP makes it easy to share the codes with the research community. A simulated hyperspectral image cube covering an area of 100 km × 100 km requires 190 GB of disk space. GCP provides a considerable amount of storing capacity at a low cost to store such huge files and their efficient access for further processing. The hyperspectral nature of this data (5 nm bandwidth) makes it appropriate for generating a detailed vegetation map with improved accuracy for a variety of applications, including fire and forest management. Better vegetation/fuel maps are critical for effective fuel treatment, i.e. identifying areas for creating fire breaks or fire lines to check the spread of wildfires. The simulated hyperspectral image can potentially be used to extract other biophysical attributes of vegetation like chlorophyll, moisture, and nitrogen, expanding its applications to other areas of vegetation research.

5. Conclusions

In this study we developed and implemented a sophisticated workflow to generate simulated AVIRIS-NG hyperspectral image data from Sentinel-2 image at Sentinel scene scale (100 X100 km) in Google's cloud environment and tested simulation reproducibility across space and time. We employed a well-established endmember selection technique to make our method more consistent and reproducible across different geographies. By improving the simulation algorithm, we were able to turn a 4.5 GB Sentinel 2 data set into a full hyperspectral image cube of 190 GB in about 2 hrs. We developed a Random Forest image classification model using training signatures from one scene and tested the Random Forest model portability on two other scenes from different sub-ecoregions. The Random Forest vegetation classification model performed satisfactorily on these two scenes, suggesting that our model can be reliably used for improved vegetation mapping in boreal Alaska. The vegetation maps generated from simulated data were more detailed and accurate based on the available field data points than the two existing vegetation maps, LANDFIRE EVT and Alaska Vegetation and Wetland Composite (AKVWC). We assessed the mapping accuracy of the latest LANDFIRE EVT (32%) and our map product (65%) using ground survey data and observed an improvement of 33% in accuracy. This

study developed a novel approach and algorithms to produce hyperspectral image from widely available multispectral images and showed its applicability in mapping vegetation and fuel in the boreal forest of Alaska with improved accuracy, which will contribute to effective forest and fire management in Alaska.

Funding

This material is based upon work supported by the National Science Foundation and under the award OIA-1757348, the State of Alaska and the U.S. Geological Survey under Grant/Cooperative agreement No. G18AP00077.

CRediT authorship contribution statement

Anushree Badola: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft. **Santosh K. Panda:** Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing. **Dar A. Roberts:** Methodology, Supervision, Writing – review & editing. **Christine F. Waigl:** Methodology, Writing – review & editing. **Randi R. Jandt:** Writing – review & editing. **Uma S. Bhatt:** Funding acquisition, Resources, Supervision, Writing – review & editing.

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

This material is based upon work supported by the National Science Foundation under the award OIA-1757348, the State of Alaska and the U.S. Geological Survey under Grant/Cooperative agreement No. G18AP00077. A heartfelt thanks to Chris Smith, Malvika Shriwas and Brooke Kubby for assisting in fieldwork. Special thanks to Abhisek Maiti for his support with the programming part of this work. Thanks to the UAF Writing Center, especially Delcenia Cosman, for proofreading this manuscript and correcting grammatical errors. Also, we express our gratitude to the two anonymous reviewers and journal editor for their constructive feedback and suggestions that helped us improve this manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jag.2022.102891.

References

Abrams, M., Crippen, R., Fujisada, H., 2020. ASTER Global Digital Elevation Model (GDEM) and ASTER Global Water Body Dataset (ASTWBD). Remote Sens. 2020, Vol. 12, Page 1156 12, 1156. https://doi.org/10.3390/RS12071156.

Ahmad, S., Pandey, A.C., Kumar, A., Lele, N.V., 2021. Potential of hyperspectral AVIRIS-NG data for vegetation characterization, species spectral separability, and mapping. Appl. Geomatics 13 (3), 361–372.

Alaska Vegetation and Wetland Composite, 2019. Alaska Vegetation and Wetland Composite | Alaska Conservation Science Catalog [WWW Document]. URL https://accscatalog.uaa.alaska.edu/dataset/alaska-vegetation-and-wetland-composite (accessed 11.4.21).

Asner, G.P., 1998. Biophysical and Biochemical Sources of Variability in Canopy Reflectance. Remote Sens. Environ. 64, 234–253. https://doi.org/10.1016/S0034-4257(98)00014-5.

Badola, A., Padalia, H., Belgiu, M., Prabhakar, M., Verma, A., 2019. Mapping Tree Species Richness of Tropical Forest using Airborne Hyperspectral Remote Sensing.

Badola, A., Padalia, H., Belgiu, M., Verma, P.A., 2021a. Tree Species Mapping in Tropical Forests Using Hyperspectral Remote Sensing and Machine Learning 5421–5424. https://doi.org/10.1109/IGARSS47720.2021.9553549.

Badola, A., Panda, S.K., Roberts, D.A., Waigl, C.F., Bhatt, U.S., Smith, C.W., Jandt, R.R., 2021b. Hyperspectral Data Simulation (Sentinel-2 to AVIRIS-NG) for Improved Wildfire Fuel Mapping, Boreal Alaska. Remote Sens. 2021, Vol. 13, Page 1693 13, 1693. https://doi.org/10.3390/RS13091693.

Breiman, L., 2001. Random Forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324.

Clark, M.L., Roberts, D.A., Clark, D., 2005. Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. Remote Sens. Environ. 96, 375–398. https://doi.org/10.1016/j.rse.2005.03.009.

Develice, R.L., 2012. Accuracy of the LANDFIRE Alaska Existing Vegetation Map over the Chugach National Forest.

Dobrinić, D., Gašparović, M., Medak, D., 2021. Sentinel-1 and 2 Time-Series for Vegetation Mapping Using Random Forest Classification: A Case Study of Northern Croatia. Remote Sens. 2021, Vol. 13, Page 2321 13, 2321. https://doi.org/10.3390/RS13122321.

Dudley, K.L., Dennison, P.E., Roth, K.L., Roberts, D.A., Coates, A.R., 2015. A multi-temporal spectral library approach for mapping vegetation species across spatial and temporal phenological gradients. Remote Sens. Environ. 167, 121–134. https://doi.org/10.1016/J.RSE.2015.05.004.

EarthData, 2021. Earthdata [WWW Document]. URL https://earthdata.nasa.gov/ (accessed 11.3.21).

European Space Agency, 2017. Sentinel-2 Spectral Response Functions (S2-SRF) - Sentinel-2 MSI Document Library - User Guides - Sentinel Online [WWW Document]. URL https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi/document-library/-/asset_publisher/Wk0TKajiISaR/content/sentinel-2a-spectral-responses (accessed 11.23.20).

European Space Agency, 2014. Copernicus Open Access Hub [WWW Document]. URL (accessed 11.23.20).

Exelis Visual Information Solutions version 5.3, 2010. Exelis Visual Information Solutions.

Fernandez-Manso, A., Quintano, C., Roberts, D.A., 2016. Burn severity influence on post-fire vegetation cover resilience from Landsat MESMA fraction images time series in Mediterranean forest ecosystems. Remote Sens. Environ. 184, 112–123. https://doi.org/10.1016/J.RSE.2016.06.015.

Gillies, S., others, 2013. Rasterio: geospatial raster I/O for {Python} programmers.

Google, 2022. Google Cloud documentation | Documentation [WWW Document]. URL https://cloud.google.com/docs (accessed 1.2.22).

Govender, M., Chetty, K., Naiken, V., Bulcock, H., 2019. A comparison of satellite hyperspectral and multispectral remote sensing imagery for improved classification and mapping of vegetation. Water SA 34, 147–154. https://doi.org/10.4314/wsa.v34i2.

Grabska, E., Hostert, P., Pflugmacher, D., Ostapowicz, K., 2019. Forest Stand Species Mapping Using the Sentinel-2 Time Series. Remote Sens. 11, 1197. https://doi.org/10.3390/rs11101197.

Harris, C.R., Millman, K.J., van der Walt, S.J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N.J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M.H., Brett, M., Haldane, A., del Río, J.F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., Oliphant, T.E., 2020. Array programming with {NumPy}. Nature 585 (7825), 357–362.

Hati, J.P., Goswami, S., Samanta, S., Pramanick, N., Majumdar, S.D., Chaube, N.R., Misra, A., Hazra, S., 2021. Estimation of vegetation stress in the mangrove forest using AVIRIS-NG airborne hyperspectral data. Model. Earth Syst. Environ. 7 (3), 1877–1889.

Hsu, P.H., 2007. Feature extraction of hyperspectral images using wavelet and matching pursuit. ISPRS Journal of Photogrammetry and Remote Sensing 62 (2), 78–92. https://doi.org/10.1016/J.ISPRSJPRS.2006.12.004.

International Arctic Research Center, 2021. Alaska's Changing Wildfire Environment | IARC [WWW Document]. URL https://uaf-iarc.org/alaskas-changing-wildfire-environment/ (accessed 11.9.21).

Landfire, 2016. LANDFIRE: Existing Vegetation Type [WWW Document]. U.S. Dep. Agric. U.S. Dep, Inter http://www.landfire.gov (accessed 2.10.21.

Liu, B., Zhang, L., Zhang, X., Zhang, B., Tong, Q., 2009. Simulation of EO-1 Hyperion Data from ALI Multispectral Data Based on the Spectral Reconstruction Approach. Sensors 9, 3090–3108. https://doi.org/10.3390/s90403090.

Louis, J., Debaecker, V., Pflug, B., Main-Knorn, M., Bieniarz, J., Mueller-Wilm, U., Cadau, E., Gascon, F., 2016. SENTINEL-2 SEN2COR: L2A Processor for Users.

McHugh, M.L., 2012. Interrater reliability: the kappa statistic. Biochem. Medica 22, 276. https://doi.org/10.11613/bm.2012.031.

Meerdink, S.K., Hook, S.J., Roberts, D.A., Abbott, E.A., 2019. The ECOSTRESS spectral library version 1.0. Remote Sens. Environ. 230, 111196.

Mudele, O., Gamba, P., 2019. Mapping vegetation in urban areas using Sentinel-2. 2019 Jt. Urban Remote Sens. Event, JURSE 2019. https://doi.org/10.1109/JURSE.2019.8809019.

NASA JPL, 2018. AVIRIS-Next Generation [WWW Document]. URL https://avirisng.jpl.nasa.gov/platform.html (accessed 11.24.20).

National Park Service, 2021. Fire in Ecosystems: Boreal Forest (U.S. National Park Service) [WWW Document]. URL https://www.nps.gov/articles/000/fire-in-ecosystems-boreal-forest.htm (accessed 11.9.21).

Nowacki, G.J., Spencer, P., Fleming, M., Brock, T., Jorgenson, T., 2003. Unified Ecoregions of Alaska: 2001. Open-File Rep. https://doi.org/10.3133/OFR2002297.

Pattison, Robert; Andersen, Hans-Erik; Gray, Andrew; Schulz, Bethany; Smith, Robert J.; Jovan, Sarah, tech. coords., 2018. Forests of the Tanana Valley State Forest and Tetlin National Wildlife Refuge, Alaska: results of the 2014 pilot inventory. Gen. Tech. Rep. PNW-GTR-967. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: Machine Learning in {P}ython. J. Mach. Learn. Res. 12, 2825–2830.

Python Core Team, 2015. Python: A dynamic, open source programming language.

Reeves, M.C., Ryan, K.C., Rollins, M.G., Thompson, T.G., 2009. Spatial fuel data products of the LANDFIRE Project. Int. J. Wildland Fire 18 (3), 250.

Roberts, D. A., Halligan, K., Dennison, P., Dudley, K., Somers, B., Crabbe, A., 2018. Viper Tools User Manual, Version 2, 91 pp. [WWW Document].

Roberts, D.A., Dennison, P.E., Roth, K.L., Dudley, K., Hulley, G., 2015. Relationships between dominant plant species, fractional cover and Land Surface Temperature in a Mediterranean ecosystem. Remote Sens. Environ. 167, 152–167. https://doi.org/10.1016/J.RSE.2015.01.026.

Roberts, D.A., Gardner, M.E., Church, R., Ustin, S.L., Green, R.O., 1997. Optimum strategies for mapping vegetation using multiple-endmember spectral mixture models. https://doi.org/10.1117/12.278930 3118, 108-119. https://doi.org/10.1117/12.278930.

Roth, K.L., Dennison, P.E., Roberts, D.A., 2012. Comparing endmember selection techniques for accurate mapping of plant species and land cover using imaging spectrometer data. Remote Sens. Environ. 127, 139–152. https://doi.org/10.1016/J.RSE.2012.08.030.

Salas, E.A.L., Subburayalu, S.K., Slater, B., Zhao, K., Bhattacharya, B., Tripathy, R., Das, A., Nigam, R., Dave, R., Parekh, P., 2020. Mapping crop types in fragmented arable landscapes using AVIRIS-NG imagery and limited field data. Int. J. Image Data Fusion 11, 33–56. https://doi.org/10.1080/19479832.2019.1706646.

Schaaf, A.N., Dennison, P.E., Fryer, G.K., Roth, K.L., Roberts, D.A., 2011. Mapping Plant Functional Types at Multiple Spatial Resolutions Using Imaging Spectrometer Data. GIScience Remote Sens. 48, 324–344. https://doi.org/10.2747/1548-1603.48.3.324.

Singh, P., Srivastava, P.K., Malhi, R.K.M., Chaudhary, S.K., Verrelst, J., Bhattacharya, B. K., Raghubanshi, A.S., 2021. Denoising AVIRIS-NG data for generation of new chlorophyll indices. IEEE Sens. J. 21 (5), 6982–6989.

Smith, C.W., Panda, S.K., Bhatt, U.S., Meyer, F.J., 2021. Improved Boreal Forest Wildfire Fuel Type Mapping in Interior Alaska using AVIRIS-NG Hyperspectral data. Remote Sens. 13, 897. https://doi.org/https://doi.org/10.3390/rs13050897.

Tiwari, V., Kumar, V., Pandey, K., Ranade, R., Agrawal, S., 2016. Simulation of the hyperspectral data using Multispectral data. In: in: International Geoscience and Remote Sensing Symposium (IGARSS). Institute of Electrical and Electronics Engineers Inc, pp. 6157–6160. https://doi.org/10.1109/IGARSS.2016.7730608.

U.S. Forest Service, 2016. PNW-FIA Interior Alaska Inventory | Pacific Northwest Research Station | PNW - US Forest Service [WWW Document]. URL https://www.fs.usda.gov/pnw/projects/pnw-fia-interior-alaska-inventory (accessed 12.12.21).

U.S. Geological Survey, 2001. BLM REA YKL 2011 Ecoregions of Alaska and Neighboring Territory [WWW Document]. URL http://agdc.usgs.gov/data/usgs/erosafo/ecoreg/ (accessed 1.1.22).

Viereck, L.A., Dyrness, C.T., Batten, A.R., Wenzlick, K.J., 1992. The Alaska vegetation classification. Gen. Tech. Rep. PNW-GTR-286. Portland, OR U.S. Dep. Agric. For. Serv. Pacific Northwest Res. Station. 278 p 286. https://doi.org/10.2737/PNW-GTR-286.

Wagner, F.H., Ferreira, M.P., Sanchez, A., Hirye, M.C.M., Zortea, M., Gloor, E., Phillips, O.L., de Souza Filho, C.R., Shimabukuro, Y.E., Aragão, L.E.O.C., 2018. Individual tree crown delineation in a highly diverse tropical forest using very high resolution satellite images. ISPRS J. Photogramm. Remote Sens. 145, 362–377. https://doi.org/10.1016/j.isprsjprs.2018.09.013.

Xie, Y., Sha, Z., Yu, M., 2008. Remote sensing imagery in vegetation mapping: a review. J. Plant Ecol. 1, 9–23. https://doi.org/10.1093/jpe/rtm005.

Zhang, C., 2014. Combining hyperspectral and lidar data for vegetation mapping in the Florida everglades. Photogramm. Eng. Remote Sensing 80, 733–743. https://doi.org/10.14358/PERS.80.8.733.

Zhang, L., Fujiwara, N., Furumi, S., Muramatsu, K., Daigo, M., Zhang, L., 2007. Assessment of the universal pattern decomposition method using MODIS and ETM data. Int. J. Remote Sens. 28, 125–142. https://doi.org/10.1080/01431160600617228.

Zhang, L., Furumi, S., Muramatsu, K., Fujiwara, N., Daigo, M., Zhang, L., 2006. Sensor-independent analysis method for hyperspectral data based on the pattern decomposition method. Int. J. Remote Sens. 27, 4899–4910. https://doi.org/10.1080/01431160600702640.