



ISSN: 2455-9377

Mapping cropland types in heterogeneous arid environments using machine learning algorithms and dataset variations on Google Earth Engine

Mohammed B. Altoom^{1,3*}, Elhadi Adam¹, Colbert M. Jackson^{1,2}

¹Faculty of Science, School of Geography, Archaeology and Environmental Studies, University of the Witwatersrand, Johannesburg 2050, South Africa, ²Department of Geography, Faculty of Natural and Agricultural Sciences, University of the Free State, Bloemfontein 9300, South Africa, ³School of Agricultural, Earth and Environmental Sciences, University of KwaZulu-Natal, Westville Campus, Durban, 4000, South Africa

ABSTRACT

Monitoring crops ensures global, regional, and national food security. It entails gathering information such as the health and growth of crops and environmental conditions. This aids farmers in making well-informed decisions, enhancing productivity, and reducing environmental impact. This, in turn, leads to improved economic outcomes and long-term agricultural sustainability. In arid and semi-arid lands, effective crop monitoring is particularly critical due to the limiting factor of water availability. This study used Sentinel-2 (S2) image collection in Google Earth Engine (GEE) to mapping crop types in North Darfur State, Sudan during the 2022 growing season (July 1 to September 30, 2022), using support vector machine (SVM) and random forest (RF) classifiers. Eight vegetation indices (VIs), i.e., normalised difference vegetation index (NDVI), enhanced vegetation index (EVI), soil-adjusted vegetation index (SAVI), green-normalised difference vegetation index (GNDVI), weighted Difference vegetation index (WDVI), red edge NDVI (NDVI_{re}), ratio-vegetation index (RVI), and normalised difference infrared index (NDII) were used as additional bands. The results show that the RF models produced an overall accuracy (OA) of 90-97% with a kappa coefficient (κ) of 0.87-0.96. The SVM models reported OA and κ values in the range of 84-95% and 0.81-0.94, respectively. Producer's (PA) and user's accuracies (UA) were in the range 83-97% and 81-98%, respectively. Highest F1 scores for both classifiers were 0.98. The findings of this study, along with the derived classification maps, would enable farmers, policymakers, and other stakeholders to make well-informed decisions regarding agricultural production, land use planning, and resource management in North Darfur and arid environments.

KEYWORDS: Sentinel-2 image collection, Crop type mapping, North Darfur State, Arid lands, Google Earth Engine

Received: July 11, 2024
Revised: December 04, 2024
Accepted: December 12, 2024
Published: February 13, 2025

***Corresponding author:**
Mohammed B. Altoom
E-mail: mohbasharadam@gmail.com

INTRODUCTION

Accurate mapping and monitoring of agricultural land at the crop type level is crucial in precision agriculture, especially in regions facing rapid climate variability (Brinkhoff *et al.*, 2019; Chabalala *et al.*, 2022). Lack of adequate data due to insufficient resources and difficulties with access to some areas, probably because of security concerns and other obstacles, means that traditional approaches to collecting agricultural data are challenging (Eklund *et al.*, 2017; Qader *et al.*, 2021). Furthermore, Crop yield assessment based on conventional data collection techniques such as crop cut data, expert scouting estimates, and field surveys lacks reliability because they are resource-intensive, time consuming, and susceptible to errors (Ge *et al.*, 2016; Sinha *et al.*, 2022). Therefore, embracing such an approach may introduce

uncertainties as yields can display significant spatial variations within farmland (Lobell *et al.*, 2019; Paliwal & Jain, 2020). As the demand for addressing food insecurity and enhancing livelihoods grows, the significance of dependable farm-level yield estimates derived from crop type maps is increasingly paramount (Sapkota *et al.*, 2016; Karst *et al.*, 2020).

To address the ongoing and growing demands for future food needs, it is vital to adopt agricultural intelligence technologies that are accurate, timely, reliable, cost-effective, and inclusive (Löw *et al.*, 2013, Ouzemou *et al.*, 2018). Such technology will facilitate improved management of the agricultural landscape and accurate mapping of crops (Defourny *et al.*, 2019; Mazarire *et al.*, 2020). Satellite remote sensing (RS) has become a crucial and efficient technology, allowing for spatiotemporal

Copyright: © The authors. This article is open access and licensed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted, use, distribution and reproduction in any medium, or format for any purpose, even commercially provided the work is properly cited. Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made.

monitoring of crops. Remote sensing (RS) helps identify crop types, evaluate actual distribution and extent of damage, assess nutrient status, and predict yields (Asgarian *et al.*, 2016; Cui *et al.*, 2022). Furthermore, RS saves time and conserve resources in agricultural management (Asgarian *et al.*, 2016; Kosari *et al.*, 2020; Felegari *et al.*, 2021; Cui *et al.*, 2022). Remote sensing (RS) offers the advantage of monitoring vast spatial expanses (Gourlay *et al.*, 2017; Böhler *et al.*, 2020; Chabalala *et al.*, 2023). Nevertheless, monitoring crops in heterogeneous arid environments faces frequent hindrances due to challenges such as limited input datasets, high soil background reflectance, the coexistence of green and senescent grasses, and the prevalence of shrubs in grasslands, which share similar reflective characteristics to cropland (Ozdogan, 2010; Tran *et al.*, 2022; Qader *et al.*, 2023). Recent improvements in temporal, spatial, and spectral resolutions, exemplified by sensor datasets like those from the Sentinel-2 Multispectral Imager (MSI), provide new opportunities for monitoring smallholder farms. The global coverage and free access to such datasets further contribute to their utility. The Sentinel-2 MSI sensor comprises shorter revisit periods to accommodate the temporal dynamics of crop growth (Malenovský *et al.*, 2012). The high spatial resolution across a broad range of spectral channels facilitates the discrimination of intricate vegetation spectral signatures (Defourny *et al.*, 2019). Furthermore, it boasts high radiometric and spectral resolution, enabling the identification of detailed vegetation properties.

Vegetation indices (VIs) serve as valuable tools in precision agriculture, such as mapping crop types for policy makers to make well-informed decisions (Thenkabail *et al.*, 2000). Due to the numerous light, environmental, and biological variables in agriculture, researchers have devised various vegetation indices (Rouse *et al.*, 1974; Xue & Su, 2017). Selecting the appropriate VI is crucial for achieving accurate results, and this choice depends on factors such as sensor type, environmental conditions, crop type, and growth stage (Hatfield *et al.*, 2019; Kang & Özdoğan, 2019). Therefore, in crop mapping, gaining insight into the inherent strengths and limitations of commonly used indices aids in selecting the most suitable one (Delavarpour *et al.*, 2021). The RGB sensors exclusively generate VIs based on visible light, capturing what is perceptible to the human eye with subtle variations (Xue & Su, 2017). On the other hand, multispectral sensors, such as Sentinel-2 (S2), provide VI choices derived from invisible infrared wavelengths, thereby uncovering details that are not discernible to the human eye (Kganyago *et al.*, 2021).

The development of machine learning (ML) algorithms presents an intriguing opportunity to understand how classifiers operate within heterogeneous agricultural landscapes (Mazarire *et al.*, 2020). Therefore, the choice of classification method is equally crucial for successful crop type mapping besides selecting the appropriate RS data (Sothe *et al.*, 2017). Different ML algorithms such as Random Forest, Support Vector Machines (SVM), Decision Trees, Gradient Boosting, and Convolutional Neural Networks (CNNs) have been employed for mapping crop types using satellite imagery *e.g.*, (Mahalanobis *et al.*, 1996; Friedman, 2002; Qian *et al.*, 2014; Jony *et al.*, 2018; Camargo *et al.*, 2019; Loukika *et al.*, 2021; Nitze *et al.*, 2012; Rodriguez-Galiano *et al.*, 2012).

Nonparametric classifiers are deemed appropriate and superior due to their capacity to disregard the dataset's assumption of a normal distribution, and they do not necessitate statistical parameters when distinguishing between image classes (Adam *et al.*, 2014; Inglada *et al.*, 2015; Sothe *et al.*, 2017; Ouzemou *et al.*, 2018). Hence, contemporary algorithms are preferred as they overcome the limitations of traditional ones such as the Maximum Likelihood Classifier (MLC) and Minimum Distance Classifier (MDC), by effectively synthesizing classification functions that can handle both discrete and continuous datasets (Sothe *et al.*, 2017). Machine learning (ML) algorithms are also robust to noise, ensuring that they are not constrained by assumptions related to data distributions (Löw *et al.*, 2013). Moreover, a gap exists in identifying which machine learning algorithms are most suitable for the distinctive classification of crop types (Mazarire *et al.*, 2020).

While the acquisition and availability of satellite imagery data have seen significant improvements, traditional hardware and software methods remain time-consuming, mainly when undertaking image processing on a larger scale (Gorelick *et al.*, 2017; Kumar & Mutanga, 2018; Mashala *et al.*, 2023). Therefore, massive challenges exist in RS data download and processing efficiency on large scale. Using high-resolution images for identifying crop types across extensive regions is expensive, demanding substantial computation power for image processing and significant storage capacities (Shaharum *et al.*, 2020; Luo *et al.*, 2021; Loukika *et al.*, 2021; Sujud *et al.*, 2021). To address these challenges effectively, Google developed the Google Earth Engine (GEE), a free cloud computing and storage platform, which enables high-speed analysis using advanced processing techniques (Sonobe *et al.*, 2014; Amani *et al.*, 2020). Thus, GEE enhances researchers' and practitioners' accessibility to computationally demanding analyses, particularly those with limited access to advanced computational resources. Google Earth Engine (GEE) incorporates freely accessible satellite imagery from Sentinel, MODIS, and Landsat, among other sources. In GEE, client libraries are created in JavaScript to handle the code editing and built-in ML algorithms (Shelestov *et al.*, 2017; Pimple *et al.*, 2018).

Like most developing countries in the Sahel region, Sudan faces multiple challenges, such as high poverty, a decline in agricultural productivity and other socioeconomic challenges and environmental degradation (Mirzabaev *et al.*, 2021). Nevertheless, agriculture constitutes approximately one-third of the GDP and is the predominant source of employment, with over 75% of the labour force engaged in agricultural and related activities (Osman, 2017). Sorghum, millet, groundnuts, wheat, and sugar cane are grown in Sudan (Ibrahim, 1978; Abbas & El-Hag, 2013; Altoom *et al.*, 2023).

In North Darfur, employing RS technology in agriculture poses specific challenges, *e.g.*, the communal landscape is primarily marked by fragmented, small parcel-sized fields with varying crop types within a pixel, contributing to highly heterogeneous crop cover (Mohammed *et al.*, 2018; Altoom *et al.*, 2023). This limitation arises from various crop types exhibiting significant variability in phenological stages within fields, encompassing early sprouting, establishment, and maturation (Marx & Loboda, 2013; Altoom *et al.*, 2023). Consequently, these variations may

result in identical spectral signatures at specific points in their development (Asgarian *et al.*, 2016). This study utilised high-resolution Sentinel-2 optical satellite dataset in GEE, to map crops in North Darfur using random forest (RF) and support vector machine (SVM) classification algorithms. Also, eight VIs, i.e., normalised difference vegetation index (NDVI), enhanced vegetation index (EVI), soil-adjusted vegetation index (SAVI), green-normalised difference vegetation index (GNDVI), weighted difference vegetation index (WDVI), red-edge NDVI (NDVI_{re}), ratio-vegetation index (RVI), and normalised difference infrared index (NDII) were extracted from the S2 image collection.

MATERIALS AND METHODS

Study Area

The North Darfur State is a vast region located in western Sudan at latitude 11° 45' 49" - 20° 00' 30" N and longitude 22° 46'

47" - 27° 29' 47" E, occupying more than half of the Darfur region (Figure 1). The highest point is Marrah Mountain (3,000 m above sea level) (Ibrahim, 1984; Altoom *et al.*, 2023). Locally, the state shares borders with Northern, North, and South Kordofan and internationally with Libya and Chad. North Darfur State comprises eleven districts, and Al Fashir is the capital city.

The inter-tropical convergence zone's (ITCZ) seasonal movement and the southwest monsoon's northerly movement determine the rainfall pattern and amount in the arid and semi-arid North Darfur state. Rain falls in the summer, i.e., from June to October, with an average annual rainfall range of 152 mm in the north and 540 mm in the south, with wettest months being July and August, where 80% of rainfall occurs (Kevane & Gray, 2008). The average annual temperature ranges from 42 °C in the hot months of April and May to 11 °C in the cold months of January and December (Mohammed *et al.*, 2018). Rainfed agriculture is the main economic activity for 85% of

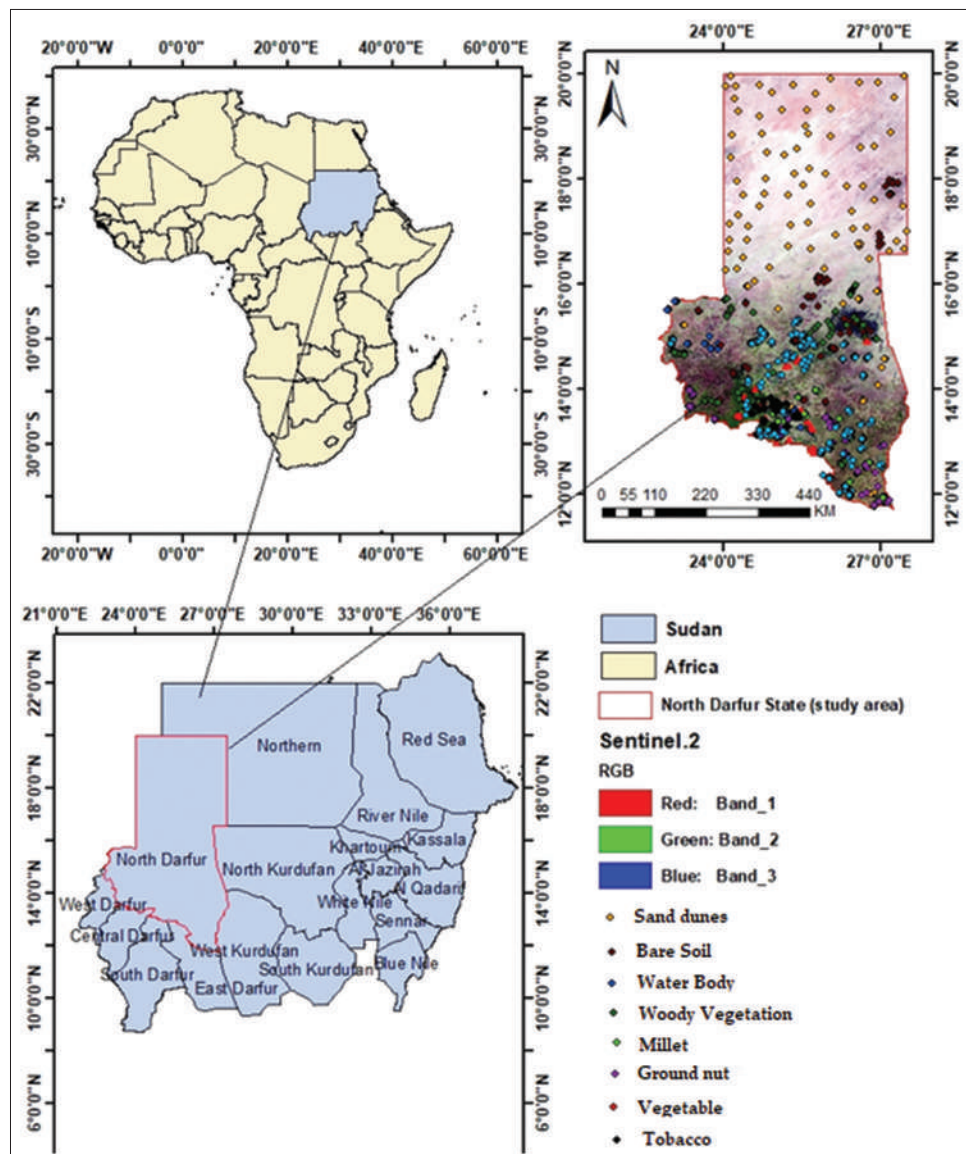


Figure 1: Geographical location of the study area in Sudan shown using Sentinel-2 (S2) false colour composite (RGB:123)

the population in the area; millet and groundnuts are grown mainly in goz lands, while tomato, onion, and okra are grown in wadi lands (Osman & Cohen, 2014). These crops are significant sources of food security, and their productivity depends on land fertility and rainfall occurrence (Osman, 2012).

Data Acquisition and Pre-processing

Several Earth Observation Data (EOD) from the previous three decades, encompassing satellite images from popular platforms such as MODIS, Landsat, and Sentinel, as well as other geographical data, are stored in the cloud-based Google Earth Engine (GEE) environment (Loukika *et al.*, 2021). These data can be accessed in the GEE cloud database through <https://earthengine.google.com/>. In this study, Sentinel-2 (S2) multispectral (MSI) Level-1C images were imported into the GEE platform and processed using the built-in `ee.ImageCollection("COPERNICUS/S2_SR")` command. The S2 satellites were launched by the European Commission (EC) and the European Space Agency (ESA) on 23 June 2015 (Sentinel-2A) and 7 March 2017 (Sentinel-2B) with a revisit frequency of 5 days (Inoue *et al.*, 2020; Xie & Niculescu, 2022). The two satellites provide continually open-access multispectral data, with both high spatial, spectral, and temporal (10 days from one satellite and 5 days by both satellites) resolutions (Table 1). The S2 bands used in this study are blue (B2), green (B3), red (B4), vegetation red edge (B5, B6, B7, and B8A), near-infrared (B8), shortwave infrared (B11), and shortwave infrared (B12). Top-of-Atmospheric (TOA)-corrected S2 (MSI-1C) data from 01/07/2022 to 30/09/2022 (a short summer marked by occasional rain and cooler daytime temperatures compared to the dry season from October to May) were obtained using the GEE JavaScript environment and utilized to map crops in this study. The results, therefore, reflect cropland conditions during this specific seasonal window. More detail about spectral bands presented in Table 1.

This study applied radiometric and geometric corrections, including orthorectification and spatial registration to a global reference system with sub-pixel accuracy for the Sentinel-2 MSI Level-1C TOA reflectance images. Because optical remote sensing (RS) images are greatly affected by cloud conditions, pixels heavily affected by dense and cirrus clouds were removed from all Sentinel-2 MSI images by applying the QA60 Quality Assessment band. All 20 m resolution bands were resampled to 10 m resolution bands by using the nearest neighbour method. This resampling algorithm has been widely used, due to its easy implementation and spectral information conservation (Roy & Dikshit, 1994). The 'filter' function was applied to filter and mosaic images. The date parameters (start and end) were set to cover the period from July 1 to September 30, 2022. The 'filter.Bound()' function was utilized to crop filtered images to match the boundaries of the study area. The resulting stacked images were normalized to account for illumination variations and reduce clouds presence. Composite imagery was generated using the median reducer to produce cloud-free composites (Abubakar *et al.*, 2023). A median value is assigned to each pixel for the entire stack of images, resulting in a single image for

the entire image collection. According to the local agricultural calendar, the growing season of crops is four months, i.e., June to September; this is considered the rain season, and it offers few cloud-free days for acquiring optical remotely sensed data. Therefore, Sentinel-2 images with minimal cloud cover (<10%) were selected for the study area. Finally, to ensure quality results, image enhancement and smoothing techniques were applied. The image acquisition dates correspond to the sprouting and/or harvesting time of crops in order to make discrimination between them easier (Mazarire *et al.*, 2020; Abubakar *et al.*, 2023).

The study selected S2 images because its bands are more suitable for monitoring vegetation, wavelength sensitivity to chlorophyll content and phenological state and also its finer spatial resolution compared to other satellite images (Sánchez-Espinosa & Schröder, 2019). Eight most widely used vegetation indices (VIs), i.e., normalised difference vegetation index (NDVI) (Rouse *et al.*, 1974), enhanced vegetation index (EVI) (Huete *et al.*, 2002; Praticò *et al.*, 2021), and soil-adjusted vegetation index (SAVI) (Huete, 1988) were computed from the S2 dataset. Others were the green normalised difference vegetation index (GNDVI) (Gitelson *et al.*, 1996), weighted difference vegetation index (WDVI) (Richardson & Wiegand, 1977), red edge NDVI (NDVI_{re}) (Frampton *et al.*, 2013), ratio-vegetation index (RVI) (Jordan, 1969), and normalised difference infrared index (NDII) (Kimes *et al.*, 1981). Table 2 shows the equations used to calculate the VIs adopted in this work and computed in GEE.

Acquisition of Ground Control Data

To improve the quality of classification results, the choice of validation and training points is a critical step (Praticò *et al.*, 2021). In this study, the target crops were millet (MI), groundnut (GN), vegetables (VE) and tobacco (TO). Other land cover classes include dunes (SD), bare soil (BS), water body (WB), woody vegetation (WV), and grassland (GL). This study collected ground control data of the different crops (Table 3) in the field in July-September 2022 using a hand-held Global Positioning System (eTrex® 20 GPS Receiver; Garmin, Olathe, KS, USA). Because North Darfur is expansive, the remaining sample data were generated via on-screen digitization on Google earth and Sentinel-2 MSI images, closely guided by visual interpretation, knowledge of the study area, and expert opinion (Figure 1). A total of 2 824 samples were shared between training and testing datasets, i.e., 70% (1977) and 30% (847), respectively. Mean reflectance curves were constructed from Sentinel-2 dataset using reference ground control points (Table 2).

Crop Type Mapping

To map crop types, five pixel-oriented classification models based on RF and SVM algorithms were trained and evaluated using the ground truth data collected in the field. The five data subsets extracted from the S2 MSI dataset are: (1) the visible and near-infrared (VNIR) bands, (2) visible and shortwave

Table 1: Properties of Sentinel-2 (S2) images used to map crop types in North Darfur State, Sudan

Bands	Description	Abbreviation	Resolution (m)	Central Wavelength (nm)
Band 1	Coastal aerosol	C	60	443.9 (S2A)/442.3 (S2B)
Band 2	Blue	B	10	496.6 (S2A)/492.1 (S2B)
Band 3	Green	G	10	560 (S2A)/559 (S2B)
Band 4	Red	R	10	664.5 (S2A)/665 (S2B)
Band 5	Vegetation Red Edge	VRE	20	703.9 (S2A)/703.8 (S2B)
Band 6	Vegetation Red Edge	VRE	20	740.2 (S2A)/739.1 (S2B)
Band 7	Vegetation Red Edge	VRE	20	782.5 (S2A)/779.7 (S2B)
Band 8	Near-infrared	NIR	10	835.1 (S2A)/833 (S2B)
Band 8A	Vegetation Red Edge	VRE	20	864.8 (S2A)/864 (S2B)
Band 9	Water vapor	W	60	945 (S2A)/943.2 (S2B)
Band 10	Shortwave-Cirrus	SWIR-C	60	1373.5 (S2A)/1376.9 (S2B)
Band 11	Shortwave Infrared	SWIR	20	1613.7 (S2A)/1610.4 (S2B)
Band 12	Shortwave Infrared	SWIR	20	2202.4 (S2A)/2185.7 (S2B)

Table 2: Vegetation indices (VIs) adopted in this work and computed in the Google Earth Engine (GEE); NIR near-infrared, SWIR shortwave infrared

Vegetation index	Abbreviation	Formula	Equation #	Reference
Normalised difference vegetation index	NDVI	$\frac{NIR - Red}{NIR + Red}$	(1)	Rouse <i>et al.</i> , 1974
Enhanced vegetation index	EVI	$\frac{2.5 \times (NIR - Red)}{NIR + 6.0 \times Red - 7.5 \times Blue + 1}$	(2)	Huete <i>et al.</i> , 2002
Soil-adjusted vegetation index	SAVI	$\frac{1.5 \times ((NIR - Red))}{((NIR + Red + 1.5))}$	(3)	Huete, 1988
Green normalised difference vegetation index	GNDVI	$\frac{NIR - Green}{NIR + Green}$	(4)	Gitelson <i>et al.</i> , 1996
Weighted difference vegetation index	WDVI	$\frac{NIR}{Green}$	(5)	Richardson & Wiegand, 1977
Rededge NDVI	NDVI _{re}	$\frac{Rededge - Red}{Rededge + Red}$	(6)	Frampton <i>et al.</i> , 2013
Ratio-vegetation index	RVI	$\frac{NIR}{Red}$	(7)	Jordan, 1969
Normalised difference infrared index	NDII	$\frac{NIR - SWIR1}{NIR + SWIR1}$	(8)	Kimes <i>et al.</i> , 1981



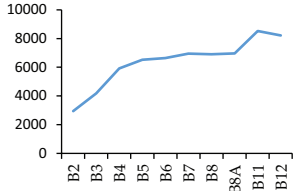


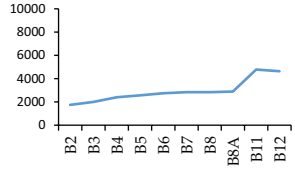


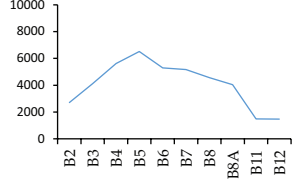


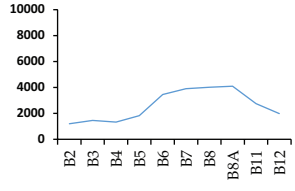


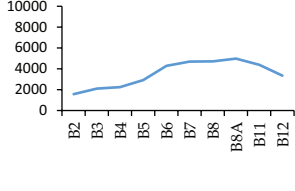




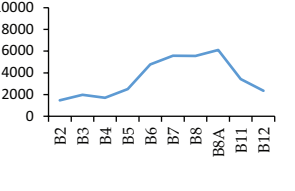
infrared (VSWIR) bands, (3) ten S2 bands, i.e., blue (band 2), green (band 3), red (band 4), vegetation red edge (bands 5, 6, 7, and B8A), near-infrared (band 8), shortwave infrared (band 11), and shortwave infrared (band 12), (4) eight vegetation indices (VIs), and (5) ten S2 bands + eight vegetation indices. The five models were later compared for crop type mapping in North Darfur State.

Image Classification

Two supervised machine learning (ML) algorithms available in the GEE platform were selected based on their different classification mechanisms, resulting in different nonlinear or linear boundaries for the same data and performed in GEE (Gorelick *et al.*, 2017). Random Forest (RF) has been one of the ensemble learning techniques that built numerous classifiers and has been proven to improve classification accuracy. RF employs a bagging (bootstrap aggregation) operation where a number of trees (*ntree*) are constructed based on a random subset of



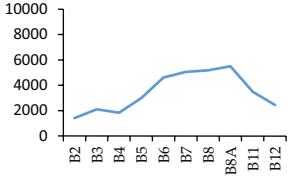


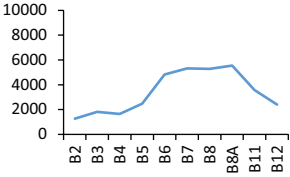
samples derived from the training data (Foody, 2002). Each tree is independently grown to maximum size based on a bootstrap sample from the training dataset without pruning. Each node is split using the best among a subset of input variables (*mtry*) (Breiman & Ihaka, 1984). Furthermore, the RF builds trees using randomly replaced bootstrapped comprising two thirds of the training samples, decreasing the variance in classification error (Genuer *et al.*, 2015). The RF has low sensitivity overtraining or noise and a high ability to handle high-dimensional data and its capacity to determine variable importance (Rodriguez-Galiano *et al.*, 2012; Xie & Niculescu, 2022). The support vector machine (SVM) classifier is a supervised classification algorithm used to solve classification and regression problems based on a statistical learning framework (Loukika *et al.*, 2021). It uses a hyperplane to classify the data points distinctly, and there are many possible methods for the hyperplane to separate the support vectors in which the key goal of SVMs is to find the hyperplane that has the maximum margin (Maxwell *et al.*, 2018; Shaharum *et al.*, 2020). The support machine learning comprises various hyperplane parameters such as kernel type,

Table 3: Reference data for cropland sites and common land cover/land use classes. The mean reflectance curves were constructed using spectral data Sentinel-2 Multi-Spectral Instrument (MSI) dataset using reference points

Crops/Other land use land cover (LULC)	Field observation images	Earth observation (sentinel-2 images)	Spectral reflectance curves	Total sample points
Sand dunes(SD)				686
Bare Soil (BS)				346
Water body (WB)				110
Woody vegetation (WV)				480
Millet (MI)				358
Groundnuts (GN)				111
Vegetables (VE)				112

(Contd...)

Table 3: (Continued)

Crops/Other land use land cover (LULC)	Field observation images	Earth observation (sentinel-2 images)	Spectral reflectance curves	Total sample points
Tobacco (TO)				95
Grassland (GL)				526

penalty value (C) and gamma (γ). These hyperparameters can be tuned and adjusted to improve the performance of SVM in classification images.

Accuracy Assessment

Confusion matrices were generated using 30% of ground reference data, from which kappa coefficient (κ), overall accuracy (OA), user accuracy (UA), producer accuracy (PA) and F1 score were computed to evaluate the performance of the models (Table 4). Overall accuracy (OA) is defined as the ratio of the total number of correctly classified pixels to the total number of pixels. The κ presented the agreements between classification and truth values, while the PA reflects the conditional probability that a specific location on the classification map output is consistent with any random sample in the test dataset. User accuracy (UA) consists of selecting a random sample with the same conditional probability as the actual type of ground from the classification maps (Rwanga & Ndambuki, 2017). The F1 score is an important metric that balances the difference between PA and UA for each class by formulating a harmonic mean.

RESULTS

Crop Type Classification Maps

The classification maps of crops in the study area were derived from optical Sentinel-2 (S2) data using random forest (RF) and support vector machine (SVM) in Google Earth Engine (GEE) environment (Figure 2). However, the results from the two classifiers show that the spatial distribution of crops across the study area is depicted differently by different models; there are visually noticeable differences among the maps.

The estimated area coverage for specific crop types in the classified maps derived from various datasets was assessed using the RF and SVM classifier algorithms. The corresponding

results are presented in Tables 5 and 6, respectively. Notably, there are differences in the area covered by each specific crop. For example, millet covered 1 359 488 ha using ten Sentinel-2 bands (bands 2, 3, 4, and 8 represent blue, green, red, and near-infrared bands, respectively, while bands 5, 6, 7, and 8A are vegetation red edge bands, and bands 11 and 12 are shortwave), but 1 426 106 ha by VSWIR (visible and shortwave infrared) bands (Table 5). Groundnuts occupied 249 525 ha, as shown by a combination of ten S2 bands and eight vegetation indices (VIs), but the VIs alone indicate that groundnuts covered 576 358 ha. Similarly, tobacco had varied spatial distribution in the study area; a combination of ten S2 bands and eight VIs showed tobacco covered 53 069 ha, while VIs alone showed 186 006 ha.

The SVM classifier showed that millet covered of 1 277 434 ha using ten S2 bands. However, VSWIR bands showed it was 1 189 634 ha. Additionally, the groundnut area was estimated as 401 928 ha by ten S2 bands, while the ten S2 bands alone showed 311 928 ha. The area covered by vegetables and tobacco was in the range of 160587-254763 ha and 48442-104417 ha, respectively, as indicated by the various datasets.

Classification Accuracy

The performance of the various data sets classified using RF and SVM algorithms was compared and presented in Tables 7, 8, and 9. The RF classifier obtained the best accuracy score of 97.1% using VNIR (visible and near-infrared) bands (Table 7). The VIs composite performed the least as compared to the other datasets. Both producer accuracy (PA) and user accuracy (UA) ranged from 83-98% (Table 9). Generally, VIs scored relatively lower PA and UA values than S2 multispectral bands. Tobacco and groundnuts reported the lowest PA and UA values. Table 10 shows the F1 scores used to assess class-wise performance for RF and SVM classifiers using five different datasets, i.e., VNIR bands, VSWIR bands, ten S2 multispectral bands, eight VIs, and ten S2 bands + eight VIs. Groundnuts

Table 4: Confusion matrix equations used to calculate kappa (κ), overall accuracy (OA), producer accuracy (PA), user accuracy (UA), and F1 score

Accuracy measure	Abbreviation	Formula	Equation #
Overall accuracy	OA	$\left(\frac{\sum_{i=1}^n P_{ii}}{N}\right) \times 100$	(9)
Kappa	κ	$\frac{N \sum_{i=1}^n P_{ii} - \sum_{i=1}^n (P_i \times P + i)}{N^2 - \sum_{i=1}^n (P_i \times P + i)}$	(10)
User accuracy	UA	$\left(\frac{P_{ii}}{P_i +}\right) \times 100$	(11)
Producer accuracy	PA	$\left(\frac{P_{ii}}{P + i}\right) \times 100$	(12)
F1-score	F1-score	$\left(\frac{AU \times PA}{AU + PA}\right) \times 2$	(13)

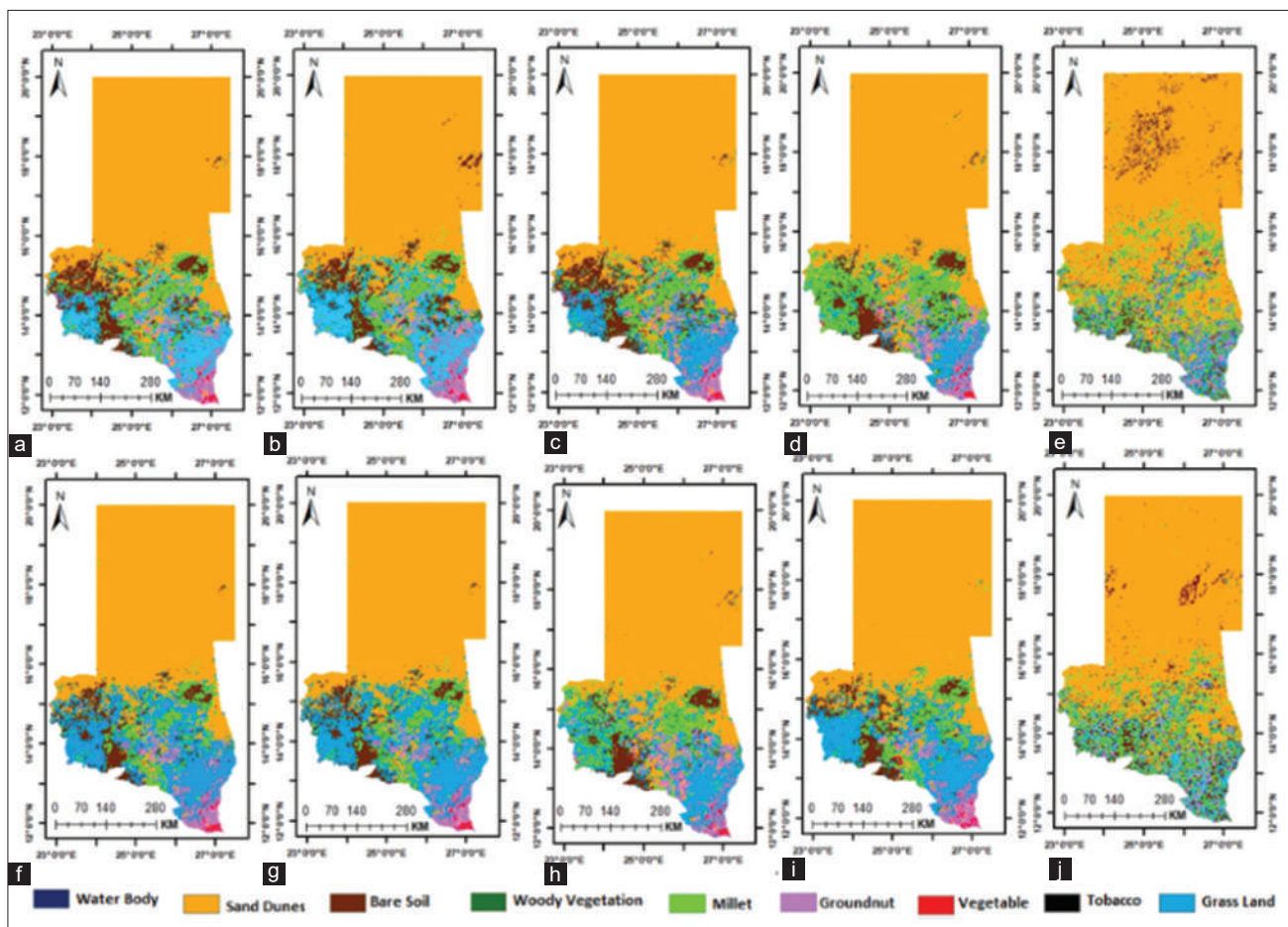


Figure 2: Classification results of spatial distribution of Crops based on the following datasets: a) RF-10 Sentinel-2 (S2) bands, b) RF-10 S2 bands + 8 vegetation indices (VIs), c) RF-visible + near-infrared (VNIR) bands, d) RF-visible + shortwave infrared (SWIR) bands, e) RF-8 Vis, f) SVM-10 S2 bands, g) SVM-10 S2 bands + 8 VIs, h) SVM-VNIR bands, i) SVM-VSWIR bands, j) SVM-8 VIs for 2022 growing season in North Darfur, Sudan. The 10 S2 bands used- B2, B3, B4, B5, B6, B7, B8, B8A, B11, B12; 8 VIs used- normalised difference vegetation index (NDVI), enhanced vegetation index (EVI), soil-adjusted vegetation index (SAVI), green-normalised difference vegetation index (GNDVI), weighted Difference vegetation index (WDVI), red edge NDVI (NDVire), ratio-vegetation index (RVI), and normalised difference infrared index (NDII); VNIR-visible and near-infrared; VSWIR-visible and shortwave

reported the highest F1 score, i.e., 0.98 from the ten S2 multispectral bands. Again, VIs performed poorly compared to the optical datasets.

However, Table 8 results show that the SVM classifier’s best overall accuracy (OA) was 95%, achieved by combining ten S2 bands and eight vegetation indices. Like with the RF classifier,

Table 5: Total estimated area (in ha) of crop types in the classified maps derived from various datasets using random forest (RF) classifier; 10 Sentinel-2 (S2) bands used - B2, B3, B4, B5, B6, B7, B8, B8A, B11, B12; 8 vegetation indices (VIs) used - normalised difference vegetation index (NDVI), enhanced vegetation index (EVI), soil-adjusted vegetation index (SAVI), green-normalised difference vegetation index (GNDVI), weighted Difference vegetation index (WDVI), red edge NDVI (NDVI_{re}), ratio-vegetation index (RVI), and normalised difference infrared index (NDII); VNIR-visible and near-infrared; VSWIR-visible and shortwave

Class	10 S2 bands	Percent (%)	10 S2 bands+8 VIs	Percent (%)	VNIR bands	Percent (%)	VSWIR bands	Percent (%)	8 VIs	Percent (%)
Sand dunes	14 644 714	63.40	14 348 403	62.12	14 543 447	62.96	14 443 447	62.53	13 896 840	60.16
Bare soil	1 929 655	8.35	1 915 642	8.29	1 561 230	6.76	1 661 230	7.19	1 826 934	7.91
Water bodies	320 097	1.38	401 163	1.73	158 517	0.69	158 517	0.68	320 167	1.39
Woody vegetation	1 609 335	6.96	1 408 074	6.096	2 354 865	10.19	2 254 865	9.76	2 191 030	9.49
Millet	1 359 488	5.88	1 395 427	6.041	1 326 106	5.74	1 426 106	6.17	1 412 910	6.12
Ground nuts	247 465	1.07	249 525	1.08	275 805	1.19	265 805	1.15	576 358	2.49
Vegetables	218 062	0.94	206 000	0.89	200 693	0.87	220 693	0.96	307 879	1.33
Tobacco	94 736	0.41	53 069	0.23	171 287	0.74	161 287	0.69	186 006	0.81
Grassland	2 672 786	11.57	3 119 035	13.50	2 504 388	10.84	2 504 388	10.84	2 378 214	10.29
Total	23 096 338	100.00	23 096 338	100.00	23 096 338	100.00	23 096 338	100.00	23 096 338	100.00

Table 6: Total estimated area (in ha) of crop types in the classified maps derived from various datasets using support vector machine (SVM) classifier: 10 Sentinel-2 (S2) bands used - B2, B3, B4, B5, B6, B7, B8, B8A, B11, B12; 8 vegetation indices (VIs) used - normalised difference vegetation index (NDVI), enhanced vegetation index (EVI), soil-adjusted vegetation index (SAVI), green-normalised difference vegetation index (GNDVI), weighted Difference vegetation index (WDVI), red edge NDVI (NDVI_{re}), ratio-vegetation index (RVI), and normalised difference infrared index (NDII); VNIR-visible and near-infrared; VSWIR-visible and shortwave

Class	10 S2 bands	Percent (%)	10 S2 bands+8 VIs	Percent (%)	VNIR bands	Percent (%)	VSWIR bands	Percent (%)	8 VIs	Percent (%)
Sand dunes	14 189 862	61.44	14 289 862	61.87	14 298 619	61.91	14 513 962	62.84	14 448 059	62.56
Bare soil	1 602 201	6.94	1 302 201	5.64	1 329 330	5.76	1 513 054	6.55	1 842 476	7.98
Water bodies	235 274	1.02	234 274	1.01	193 532	0.84	228 867	0.99	373 776	1.62
Woody vegetation	1 427 283	6.18	1 428 283	6.18	1 992 935	8.63	1 540 794	6.67	1 820 599	7.88
Millet	1 277 434	5.53	1 487 434	6.44	1 546 018	6.69	1 189 634	5.15	1 367 438	5.92
Groundnuts	311 928	1.35	401 928	1.74	253 456	1.10	267 919	1.16	214 899	0.93
Vegetables	182 557	0.79	192 557	0.83	254 763	1.10	203 773	0.88	160 587	0.69
Tobacco	58 442	0.25	48 442	0.21	104 417	0.45	61 512	0.27	98 635	0.42
Grassland	3 811 357	16.50	3 711 357	16.07	3 123 268	13.52	3 576 823	15.49	2 769 869	11.99
Total	23 096 338	100.00	23 096 338	100.00	23 096 338	100.00	23 096 338	100.00	23 096 338	100.00

the VI data set performed the least, achieving an OA value of 84%. The PA was 81-94%, while UA ranged between 87% and 97% (Table 9). Generally, VIs scored relatively lower PA and UA values than S2 multispectral bands. Yet again, Tobacco and groundnuts reported the lowest PA and UA values, while the S2 multispectral bands outperformed the VIs dataset. Millet and vegetables reported the highest F1 score, i.e., 0.94, each, by VNIR bands and the ten S2 bands/eight VIs dataset, respectively (Table 10).

DISCUSSION

This study aimed to use Google Earth Engine (GEE) to compute metrics used in mapping crop types in North Darfur State, Sudan, using datasets extracted from high-resolution Sentinel-2 (S2) optical data. Two machine learning (ML) classification algorithms, i.e., random forest (RF) and support vector machine (SVM), were used. The application of ML algorithms in mapping the spatial distribution of crops has increased in the last three decades (Rao *et al.*, 2021). However, mapping crops in highly heterogeneous arid environments is difficult due to the similarity of the

reflectance value of the pixels (Ozdogan, 2010). Also, remote sensing (RS) data for agriculture monitoring faces high-cost constraints (Khanal *et al.*, 2020). Open access RS datasets and the computing power of the GEE platform provide free access to vast volumes of satellite data and facilitate powerful processing of big geo data covering large areas (Shelestov *et al.*, 2017; Pimple *et al.*, 2018; Sujud *et al.*, 2021; Abubakar *et al.*, 2023).

The successful launch of the satellite constellation of Sentinel-2A & B by the European Space Agency (ESA) has provided unprecedented high spatial, temporal and spectral resolution data for land applications and has shown their potential in diverse applications such as agriculture (Vuolo *et al.*, 2018; Defourny *et al.*, 2019; Segarra *et al.*, 2020). The visible and near-infrared (VNIR) bands of S2 represent a promising data source for mapping crop types in regions characterized by small field areas (Tran *et al.*, 2022). The S2'S VNIR and the ten S2 bands (B2, B3, B4, B5, B6, B7, B8, B8A, B11, B12) datasets reported relatively higher classification accuracies than the other datasets for the RF classifier, i.e., 97.1% and 97%, respectively (Tables 7 and 8). An integrated dataset comprising ten S2 bands

Table 7: Confusion matrices for random forest (RF) models used to map crop types in North Darfur State, Sudan for 2022 growing season; SD-sand dunes, BS-bare soil, WB-water bodies, WV-woody vegetation, MI-millet, GN-groundnuts, VE-vegetables, TO-tobacco, GL-grassland. 10 Sentinel-2 (S2) bands used - B2, B3, B4, B5, B6, B7, B8, B8A, B11, B12; 8 vegetation indices (VIs) used - normalised difference vegetation index (NDVI), enhanced vegetation index (EVI), soil-adjusted vegetation index (SAVI), green-normalised difference vegetation index (GNDVI, weighted Difference vegetation index (WDVI), red edge NDVI (NDVI_{re}), ratio-vegetation index (RVI), and normalised difference infrared index (NDII)

10 Sentinel-2 bands									
Classes	SD	BS	WB	WV	MI	GN	VE	TO	GL
SD	477	1	0	1	3	0	0	0	1
BS	2	215	0	2	1	0	0	2	2
WB	0	1	78	0	0	0	0	0	0
WV	1	6	0	332	0	0	0	0	1
MI	2	0	1	2	244	1	0	0	2
GN	0	0	0	0	1	71	0	0	0
VE	1	0	1	1	1	0	68	0	1
TO	0	0	0	1	0	0	2	58	0
GL	1	5	1	4	1	0	2	0	331
Total	484	228	81	343	251	72	72	60	338
Overall accuracy (OA) 97%									
Kappa Coefficient (κ) 0.96									
10 Sentinel-2 (S2) bands+8 vegetation indices (VIs)									
Classes	SD	BS	WB	WV	MI	GN	VE	TO	GL
SD	467	3	1	2	3	3	1	0	3
BS	8	202	2	6	0	0	0	0	6
WB	1	2	72	1	1	0	0	0	2
WV	2	9	0	324	0	0	0	1	4
MI	3	1	0	3	239	2	1	1	2
GN	0	0	0	1	5	65	0	0	1
VE	0	1	1	1	1	0	64	1	4
TO	0	2	0	4	0	0	1	53	1
GL	8	8	0	9	2	1	0	0	317
Total	489	228	76	351	251	71	67	56	340
Overall accuracy (OA) 96.5%									
Kappa Coefficient (κ) 0.94									
Visible+ near-infrared (VNIR) bands									
Classes	SD	BS	WB	WV	MI	GN	VE	TO	GL
SD	475	1	0	2	3	1	0	0	1
BS	2	214	0	3	0	0	0	1	4
WB	0	1	78	0	0	0	0	0	0
WV	3	9	0	324	2	0	1	0	1
MI	0	1	1	2	247	0	0	0	1
GN	0	0	0	0	4	68	0	0	0
VE	1	0	1	0	2	0	68	1	0
TO	0	0	1	1	0	0	1	58	0
GL	3	4	1	3	3	1	0	0	330
Total	484	230	82	335	261	70	70	60	337
Overall accuracy (OA) 97.1%									
Kappa Coefficient (κ) 0.95									
Visible+shortwave infrared (VSWIR) bands									
Classes	SD	BS	WB	WV	MI	GN	VE	TO	GL
SD	475	2	0	2	3	1	0	0	0
BS	3	212	1	3	0	0	0	2	3
WB	1	0	78	0	0	0	0	0	0
WV	1	10	0	324	0	0	0	1	4
MI	0	0	1	0	245	1	0	0	5
GN	0	0	0	0	4	67	0	0	1
VE	1	0	1	0	0	0	69	0	2
TO	0	0	0	2	0	0	2	57	0
GL	2	6	2	4	6	2	0	0	323
Total	483	230	83	335	258	71	71	60	338

(Contd...)

Table 7: (Continued)

Overall accuracy (OA) 96.5%									
Kappa Coefficient (κ) 94%									
8 vegetation indices (VIs)									
	SD	BS	WB	WV	MI	GN	VE	TO	GL
SD	460	4	0	7	4	0	1	2	5
BS	16	171	3	9	8	5	2	0	10
WB	1	0	78	0	0	0	0	0	0
WV	19	9	1	302	5	0	0	0	4
MI	5	2	1	4	234	1	0	1	4
GN	1	2	0	1	3	62	1	0	2
VE	2	1	1	1	3	0	64	1	0
TO	1	0	1	4	3	0	1	51	0
GL	11	5	2	11	10	2	3	2	299
Total	516	194	87	339	270	70	72	57	324
Overall accuracy (OA) 90%									
Kappa coefficient (κ) 0.87									

Table 8: Confusion matrices for support vector machine (SVM) models used to map crop types in North Darfur State, Sudan for 2022 growing season; SD-sand dunes, BS-bare soil, WB-water bodies, WV-woody vegetation, MI-millet, GN-groundnuts, VE-vegetables, TO-tobacco, GL-grassland. 10 Sentinel-2 (S2) bands used - B2, B3, B4, B5, B6, B7, B8, B8A, B11, B12; 8 vegetation indices (VIs) used - normalised difference vegetation index (NDVI), enhanced vegetation index (EVI), soil-adjusted vegetation index (SAVI), green-normalised difference vegetation index (GNDVI), weighted Difference vegetation index (WDVI), red edge NDVI (NDVI_{re}), ratio-vegetation index (RVI), and normalised difference infrared index (NDII)

10 Sentinel-2 bands									
Classes	SD	BS	WB	WV	MI	GN	VE	TO	GL
SD	467	3	1	2	3	3	1	0	3
BS	8	202	2	6	0	0	0	0	6
WB	1	2	72	1	1	0	0	0	2
WV	2	9	0	324	0	0	0	1	4
MI	3	1	0	3	239	2	1	1	2
GN	0	0	0	1	5	65	0	0	1
VE	0	1	1	1	1	0	64	1	4
TO	0	2	0	4	0	0	1	53	1
GL	8	8	0	9	2	1	0	0	317
Total	490	228	76	351	251	71	67	56	340
Overall accuracy (OA) 93%									
Kappa coefficient (κ) 0.92									
10 Sentinel-2 bands+8 vegetation indices									
	SD	BS	WB	WV	MI	GN	VE	TO	GL
SD	465	6	1	3	3	1	1	1	2
BS	6	207	2	0	1	4	2	1	1
WB	0	2	74	1	0	0	2	0	0
WV	1	10	3	317	1	0	3	2	3
MI	3	4	3	0	225	4	6	3	4
GN	0	1	1	0	0	63	3	0	4
VE	0	1	2	2	1	0	64	3	0
TO	0	3	0	1	0	0	0	56	1
GL	3	5	6	8	2	1	2	3	315
Total	478	239	92	332	233	73	83	69	330
Overall accuracy (OA) 95%									
Kappa coefficient (κ) 0.94									
Visible+near-infrared bands									
	SD	BS	WB	WV	MI	GN	VE	TO	GL
SD	466	2	1	3	1	2	1	0	7
BS	5	200	2	6	2	0	2	0	7

(Contd...)

Table 8: (Continued)

Overall accuracy (OA) 95%									
Kappa coefficient (κ) 0.94									
Visible+near-infrared bands									
WB	0	1	75	1	0	0	1	0	1
WV	8	9	0	320	0	0	0	1	2
MI	6	1	0	3	237	3	0	1	1
GN	2	0	0	1	2	64	0	0	3
VE	0	1	1	3	0	0	65	1	2
TO	0	2	0	3	1	0	1	54	0
GL	11	5	2	3	5	2	0	0	317
Total	498	221	81	343	248	71	70	57	340
Overall accuracy (OA) 93%									
Kappa coefficient (κ) 0.91									
Visible+shortwave infrared bands									
	SD	BS	WB	WV	MI	GN	VE	TO	GL
SD	471	2	0	1	4	2	0	0	3
BS	6	194	1	14	1	0	0	0	8
WB	1	0	73	1	3	0	0	0	1
WV	3	9	2	320	0	0	0	0	6
MI	4	1	1	5	228	4	2	1	6
GN	2	0	0	0	7	61	1	0	1
VE	1	1	0	2	0	0	64	1	4
TO	0	3	0	2	0	0	1	54	1
GL	7	7	1	7	5	2	0	0	316
Total	495	217	78	352	248	69	68	56	346
Overall accuracy (OA) 92%									
Kappa coefficient (κ) 0.90									
8 vegetation indices									
	SD	BS	WB	WV	MI	GN	VE	TO	GL
SD	460	2	2	7	6	1	1	0	4
BS	14	179	1	9	6	2	0	1	12
WB	0	0	78	0	0	0	0	1	0
WV	9	8	1	307	5	0	1	0	9
MI	3	2	1	6	236	1	1	0	2
GN	0	1	0	0	5	63	1	0	2
VE	1	2	1	2	2	1	62	1	1
TO	2	3	0	2	1	0	1	52	0
GL	9	12	5	9	10	2	2	2	294
Total	498	209	89	342	271	70	69	57	324
Overall accuracy (OA) 84%									
Kappa coefficient (κ) 0.81									

and eight vegetation indices (VIs) achieved the same overall accuracy (OA) as the VSWIR dataset, i.e., 96.5%. Therefore, VIs in the ten S2 bands/VIs dataset improved the negative effects associated with large sets of correlated variables and decreased the classification accuracies (Tables 7, 8 & 10). The eight VIs dataset reported the lowest classification accuracies reported by both RF and SVM. Hence, the classification accuracies for the VIs dataset did not achieve the highest levels, similar to reports by Kobayashi *et al.* (2020) and Zhang *et al.* (2020); but the findings of the current study agree with Pasternak and Pawluszek-Filipiak (2022). Also, SWIR (shortwave infrared) bands had the same effect because of redundant information. Hence, the utilization of additional variables does not necessarily improve classification accuracy (Pasternak & Pawluszek-Filipiak, 2022).

The results of this study are in line with previous studies conducted by (Praticò *et al.*, 2021; Rao *et al.*, 2021) whom reported that the optical images was able to map cropland types at the field-scale with high accuracy in smallholder systems. Also the study results are in agreement with Sujud *et al.* (2021) who achieved high classification accuracies using machine learning classifiers while mapping crop types using S2 data.

Overall, the RF algorithm offered the best model for mapping crops than SVM, although the gap between the two was not significant. The RF combine numerous soft linear boundaries at the surface of decision, while SVM performs well with sparse training, making it a better choice when less data are available (Shetty, 2019). Random Forest algorithm is resilient and less affected by parameters such as outliers, missing data, and overfitting tendencies, while SVM is more sensitive to

Table 9: The user and producer accuracy performance of random forest (RF) and support vector machine (SVM) using various datasets; SD-sand dunes, BS-bare soil, WB-water bodies, WV-woody vegetation, MI-millet, GN-groundnuts, VE-vegetables, TO-tobacco, GL-grassland, S2-Sentinel-2, VI-vegetation index, VNIR-visible and near-infrared, VSWIR-visible and shortwave, RF-random forest, SVM-support vector machine. 10 S2 bands used - B2, B3, B4, B5, B6, B7, B8, B8A, B11, B12; 8 VIs used - normalised difference vegetation index (NDVI), enhanced vegetation index (EVI), soil-adjusted vegetation index (SAVI), green-normalised difference vegetation index (GNDVI, weighted Difference vegetation index (WDVI), red edge NDVI (NDVIre), ratio-vegetation index (RVI), and normalised difference infrared index (NDII)

RF Producers Accuracy						RF Users Accuracy				
Class	10 S2 bands	10 S2 bands+8 VIs	VNIR bands	VSWIR bands	8 VIs	10 S2 bands	10 S2 bands+8 VIs	VNIR bands	VSWIR bands	8 VIs
SD	0.98	0.98	0.98	0.98	0.95	0.98	0.98	0.98	0.98	0.91
BS	0.96	0.94	0.95	0.95	0.80	0.94	0.92	0.93	0.93	0.86
WB	0.98	1.00	0.98	0.97	0.97	0.96	0.96	0.95	0.96	0.92
WV	0.97	0.95	0.95	0.96	0.88	0.96	0.95	0.96	0.95	0.86
MI	0.96	0.97	0.98	0.96	0.92	0.97	0.94	0.94	0.93	0.88
GN	0.98	0.93	0.94	0.91	0.83	0.98	0.94	0.97	0.97	0.88
VE	0.93	0.94	0.93	0.93	0.87	0.94	0.97	0.97	0.97	0.91
TO	0.95	0.93	0.95	0.91	0.83	0.96	0.95	0.96	0.98	0.83
GL	0.96	0.93	0.95	0.98	0.86	0.97	0.95	0.97	0.98	0.91
SVM Producers Accuracy						SVM Users Accuracy				
SD	0.96	0.97	0.96	0.97	0.95	0.95	0.99	0.93	0.95	0.89
BS	0.90	0.94	0.89	0.86	0.78	0.88	0.90	0.90	0.89	0.84
WB	0.91	0.88	0.94	0.92	0.89	0.94	0.93	0.92	0.93	0.85
WV	0.95	0.92	0.94	0.94	0.88	0.92	0.94	0.94	0.90	0.85
MI	0.94	0.91	0.94	0.90	0.89	0.95	0.94	0.95	0.91	0.89
GN	0.90	0.91	0.88	0.84	0.81	0.91	0.93	0.90	0.88	0.87
VE	0.87	0.96	0.89	0.87	0.82	0.95	0.92	0.92	0.94	0.91
TO	0.86	0.86	0.88	0.88	0.82	0.94	0.97	0.94	0.96	0.88
GL	0.91	0.96	0.91	0.91	0.83	0.93	0.91	0.93	0.91	0.91

Table 10: F1 scores class-wise performance of random forest (RF) and support vector machine (SVM) algorithms using five different datasets; SD-sand dunes, BS-bare soil, WB-water bodies, WV-woody vegetation, MI-millet, GN-groundnuts, VE-vegetables, TO-tobacco, GL-grassland, S2-Sentinel-2, VI-vegetation index, VNIR-visible and near-infrared, VSWIR-visible and shortwave. 10 S2 bands used - B2, B3, B4, B5, B6, B7, B8, B8A, B11, B12; 8 VIs used - normalised difference vegetation index (NDVI), enhanced vegetation index (EVI), soil-adjusted vegetation index (SAVI), green-normalised difference vegetation index (GNDVI, weighted Difference vegetation index (WDVI), red edge NDVI (NDVIre), ratio-vegetation index (RVI), and normalised difference infrared index (NDII)

Class	Random forest (RF)					Support vector machine (SVM)				
	10 S2 bands	10 S2 bands+8 VIs	VNIR	VSWIR	8 VIs	10 S2 bands	10 S2 bands+8 VIs	VNIR	VSWIR	8 VIs
SD	0.98	0.98	0.98	0.98	0.93	0.95	0.98	0.94	0.95	0.92
BS	0.95	0.93	0.93	0.93	0.83	0.89	0.92	0.89	0.87	0.81
WB	0.97	0.90	0.96	0.95	0.94	0.92	0.90	0.92	0.92	0.87
WV	0.96	0.95	0.95	0.96	0.87	0.93	0.93	0.94	0.91	0.86
MI	0.96	0.95	0.95	0.94	0.89	0.94	0.92	0.94	0.90	0.89
GN	0.98	0.93	0.95	0.93	0.85	0.90	0.92	0.88	0.85	0.84
VE	0.93	0.95	0.94	0.94	0.89	0.90	0.94	0.90	0.90	0.86
TO	0.95	0.94	0.95	0.93	0.83	0.90	0.91	0.90	0.91	0.85
GL	0.96	0.93	0.95	0.97	0.88	0.92	0.93	0.91	0.91	0.87

hyperparameters (Chang *et al.*, 2019). Considering the total area occupied by each crop type (Tables 5 & 6), it was observed that the two classifiers produced quite different results. Based on the classified maps, the results show that there existed confusion among classes; due to similarity of reflectance by different classes, e.g., tobacco and vegetables. Based on results presented in Tables 9 and 10, millet recorded highest producer accuracy (PA), user accuracy (UA), and F1 scores, followed by groundnuts; the reason could be due to their unique reflectance characteristics, and higher number of ground reference data,

especially for millet (groundnuts class had comparatively fewer sample data collected in the field).

CONCLUSION

Accurate spatial distribution information for crops is important for agricultural management and food security. In this research, the mapping of crop types in North Darfur State was carried out using random forest (RF) and support vector machine (SVM) classifiers with the Sentinel-2 (S2) optical dataset within the

Google Earth Engine (GEE) environment. By analysing and comparing classification performance, RF and SVM algorithms could map crops in the study area; the classifiers produced acceptable results. To enhance classification accuracy, images from key growth periods for each type of crop cultivated should be considered when designing methodologies for crop type mapping. However, data gaps in optical time series existed because of cloud contamination. Furthermore, reliable cloud and cloud shadow detection algorithms are lacking. Future research will consider optical images from alternative sources, and the incorporation of radar images, which remain unaffected by time of day or weather conditions; data integration and more sophisticated data fusion techniques are required. Additionally, the reliance on error matrices to estimate classification accuracy based on field-collected samples introduces bias in conclusions. Consequently, future research will evaluate model performance using metrics such as balanced accuracy, bias score, Matthew's Correlation Coefficient, and others.

ACKNOWLEDGEMENTS

The authors acknowledge Google Earth Engine (GEE) for providing easy access to Sentinel-2 (S2) multispectral (MSI) and processing. We also would like to thank the North Darfur State Ministry of Agriculture for providing the agricultural data.

REFERENCES

- Abbas, H. M., & El-Hag, A. M. H. (2013). Crop assessment and monitoring for sugarcane crop, Sudan (New Halfa case study) using remote sensing and GIS. *International Journal of Scientific and Research Publications*, 3(3), 1-10.
- Abubakar, G. A., Wang, K., Koko, A. F., Husseini, M. I., Shuka, K. A. M., Deng, J., & Gan, M. (2023). Mapping Maize Cropland and Land Cover in Semi-Arid Region in Northern Nigeria Using Machine Learning and Google Earth Engine. *Remote Sensing*, 15(11), 2835. <https://doi.org/10.3390/rs15112835>
- Adam, E., Mutanga, O., Odindi, J., & Abdel-Rahman, E. M. (2014). Land-use/cover classification in a heterogeneous coastal landscape using RapidEye imagery: evaluating the performance of random forest and support vector machines classifiers. *International Journal of Remote Sensing*, 35(10), 3440-3458. <https://doi.org/10.1080/01431161.2014.903435>
- Altoom, M. B., Adam, E., & Ali, K. A. (2023). Mapping and Monitoring Spatio-Temporal Patterns of Rainfed Agriculture Lands of North Darfur State, Sudan, Using Earth Observation Data. *Land*, 12(2), 307. <https://doi.org/10.3390/land12020307>
- Amani, M., Ghorbanian, A., Ahmadi, S. A., Kakooei, M., Moghimi, A., Mirmazloumi, S. M., Moghaddam, S. H. A., Mahdavi, S., Ghahremanloo, M., Parsian, S., Wu, Q., & Brisco, B. (2020). Google earth engine cloud computing platform for remote sensing big data applications: A comprehensive review. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, 5326-5350. <https://doi.org/10.1109/JSTARS.2020.3021052>
- Asgarian, A., Soffianian, A., & Pourmanafi, S. (2016). Crop type mapping in a highly fragmented and heterogeneous agricultural landscape: A case of central Iran using multi-temporal Landsat 8 imagery. *Computers and Electronics in Agriculture*, 127, 531-540. <https://doi.org/10.1016/j.compag.2016.07.019>
- Böhler, J. E., Schaepman, M. E., & Kneubühler, M. (2020). Crop separability from individual and combined airborne imaging spectroscopy and UAV multispectral data. *Remote Sensing*, 12(8), 1256. <https://doi.org/10.3390/rs12081256>
- Breiman, L., & Ihaka, R. (1984). *Nonlinear discriminant analysis via scaling and ACE*. Technical Report No. 40. Department of Statistics, University of California.
- Brinkhoff, J., Vardanega, J., & Robson, A. J. (2019). Land cover classification of nine perennial crops using sentinel-1 and-2 data. *Remote Sensing*, 12(1), 96. <https://doi.org/10.3390/rs12010096>
- Camargo, F. F., Sano, E. E., Almeida, C. M., Mura, J. C., & Almeida, T. (2019). A comparative assessment of machine-learning techniques for land use and land cover classification of the Brazilian tropical savanna using ALOS-2/PALSAR-2 polarimetric images. *Remote Sensing*, 11(13), 1600. <https://doi.org/10.3390/rs11131600>
- Chabalala, Y., Adam, E., & Ali, K. A. (2022). Machine Learning Classification of Fused Sentinel-1 and Sentinel-2 Image Data towards Mapping Fruit Plantations in Highly Heterogeneous Landscapes. *Remote Sensing*, 14(11), 2621. <https://doi.org/10.3390/rs14112621>
- Chabalala, Y., Adam, E., & Ali, K. A. (2023). Exploring the Effect of Balanced and Imbalanced Multi-Class Distribution Data and Sampling Techniques on Fruit-Tree Crop Classification Using Different Machine Learning Classifiers. *Geomatics*, 3(1), 70-92. <https://doi.org/10.3390/geomatics3010004>
- Chang, K.-T., Merghadi, A., Yunus, A. P., Pham, B. T., & Dou, J. (2019). Evaluating scale effects of topographic variables in landslide susceptibility models using GIS-based machine learning techniques. *Scientific Reports*, 9, 12296. <https://doi.org/10.1038/s41598-019-48773-2>
- Cui, J., Zhu, M., Liang, Y., Qin, G., Li, J., & Liu, Y. (2022). Land use/land cover change and their driving factors in the Yellow River Basin of Shandong Province based on google earth Engine from 2000 to 2020. *ISPRS International Journal of Geo-Information*, 11(3), 163. <https://doi.org/10.3390/ijgi11030163>
- Defourny, P., Bontemps, S., Bellemans, N., Cara, C., Dedieu, G., Guzzonato, E., Hagolle, O., Inglada, J., Nicola, L., Rabaute, T., Savinaud, M., Udroui, C., Valero, S., Bégué, A., Dejoux, J.-F., El Harti, A., Ezzahar, J., Kussul, N., Labbassi, K., Koetz, B. (2019). Near real-time agriculture monitoring at national scale at parcel resolution: Performance assessment of the Sen2-Agri automated system in various cropping systems around the world. *Remote sensing of Environment*, 221, 551-568. <https://doi.org/10.1016/j.rse.2018.11.007>
- Delavarpour, N., Koparan, C., Nowatzki, J., Bajwa, S., & Sun, X. (2021). A technical study on UAV characteristics for precision agriculture applications and associated practical challenges. *Remote Sensing*, 13(6), 1204. <https://doi.org/10.3390/rs13061204>
- Eklund, L., Abdi, A., & Islar, M. (2017). From producers to consumers: the challenges and opportunities of agricultural development in Iraq Kurdistan. *Land*, 6(2), 44. <https://doi.org/10.3390/land6020044>
- Felegari, S., Sharifi, A., Moravej, K., Amin, M., Golchin, A., Muzirafuti, A., Tariq, A., & Zhao, N. (2021). Integration of sentinel 1 and sentinel 2 satellite images for crop mapping. *Applied Sciences*, 11(21), 10104. <https://doi.org/10.3390/app112110104>
- Foody, G. M. (2002). Status of land cover classification accuracy assessment. *Remote sensing of Environment*, 80(1), 185-201. [https://doi.org/10.1016/S0034-4257\(01\)00295-4](https://doi.org/10.1016/S0034-4257(01)00295-4)
- Frampton, W. J., Dash, J., Watmough, G., & Milton, E. J. (2013). Evaluating the capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 82, 83-92. <https://doi.org/10.1016/j.isprsjprs.2013.04.007>
- Friedman, J. H. (2002). Stochastic gradient boosting. *Computational Statistics & Data Analysis*, 38(4), 367-378. [https://doi.org/10.1016/S0167-9473\(01\)00065-2](https://doi.org/10.1016/S0167-9473(01)00065-2)
- Ge, Y., Bai, G., Stoerger, V., & Schnable, J. C. (2016). Temporal dynamics of maize plant growth, water use, and leaf water content using automated high throughput RGB and hyperspectral imaging. *Computers and Electronics in Agriculture*, 127, 625-632. <https://doi.org/10.1016/j.compag.2016.07.028>
- Genuer, R., Poggi, J.-M., & Tuleau-Malot, C. (2015). VSURF: an R package for variable selection using random forests. *The R Journal*, 7(2), 19-33.
- Gitelson, A. A., Kaufman, Y. J., & Merzlyak, M. N. (1996). Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote sensing of Environment*, 58(3), 289-298. [https://doi.org/10.1016/S0034-4257\(96\)00072-7](https://doi.org/10.1016/S0034-4257(96)00072-7)
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202, 18-27. <https://doi.org/10.1016/j.rse.2017.06.031>
- Gourlay, S., Kilic, T., & Lobell, D. (2017). Could the debate be over? errors in farmer-reported production and their implications for the inverse scale-productivity relationship in Uganda. *World Bank Policy Research Working Paper*.
- Hatfield, J. L., Prueger, J. H., Sauer, T. J., Dold, C., O'Brien, P., & Wacha, K.

- (2019). Applications of vegetative indices from remote sensing to agriculture: Past and future. *Inventions*, 4(4), 71. <https://doi.org/10.3390/inventions4040071>
- Huete, A. R. (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, 25(3), 295-309. [https://doi.org/10.1016/0034-4257\(88\)90106-X](https://doi.org/10.1016/0034-4257(88)90106-X)
- Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 83(1-2), 195-213. [https://doi.org/10.1016/S0034-4257\(02\)00096-2](https://doi.org/10.1016/S0034-4257(02)00096-2)
- Ibrahim, F. (1978). Anthropogenic causes of desertification in Western Sudan. *GeoJournal*, 2, 243-254. <https://doi.org/10.1007/BF00208640>
- Ibrahim, F. N. (1984). Ecological imbalance in the Republic of the Sudan: with reference to desertification in Darfur.
- Inglada, J., Arias, M., Tardy, B., Hagolle, O., Valero, S., Morin, D., Dedieu, G., Sepulcre, G., Bontemps, S., & Defourny, P. (2015). Assessment of an operational system for crop type map production using high temporal and spatial resolution satellite optical imagery. *Remote Sensing*, 7(9), 12356-12379. <https://doi.org/10.3390/rs70912356>
- Inoue, S., Ito, A., & Yonezawa, C. (2020). Mapping Paddy fields in Japan by using a Sentinel-1 SAR time series supplemented by Sentinel-2 images on Google Earth Engine. *Remote Sensing*, 12(10), 1622. <https://doi.org/10.3390/rs12101622>
- Jony, R. I., Woodley, A., Raj, A., & Perrin, D. (2018). Ensemble classification technique for water detection in satellite images. *2018 Digital Image Computing: Techniques and Applications (DICTA)*, 2018, 1-8. <https://doi.org/10.1109/DICTA.2018.8615870>
- Jordan, C. F. (1969). Derivation of leaf-area index from quality of light on the forest floor. *Ecology*, 50(4), 663-666. <https://doi.org/10.2307/1936256>
- Kang, Y., & Özdoğan, M. (2019). Field-level crop yield mapping with Landsat using a hierarchical data assimilation approach. *Remote Sensing of Environment*, 228, 144-163. <https://doi.org/10.1016/j.rse.2019.04.005>
- Karst, I. G., Mank, I., Traoré, I., Sorgho, R., Stüeckemann, K.-J., Simboro, S., Sié, A., Franke, J., & Sauerborn, R. (2020). Estimating yields of household fields in rural subsistence farming systems to study food security in Burkina Faso. *Remote Sensing*, 12(11), 1717. <https://doi.org/10.3390/rs12111717>
- Kevane, M., & Gray, L. (2008). Darfur: rainfall and conflict. *Environmental Research Letters*, 3(3), 034006. <https://doi.org/10.1088/1748-9326/3/3/034006>
- Kganyago, M., Mhangara, P., & Adjorlolo, C. (2021). Estimating crop biophysical parameters using machine learning algorithms and Sentinel-2 imagery. *Remote Sensing*, 13(21), 4314. <https://doi.org/10.3390/rs13214314>
- Khanal, S., Kushal, K. C., Fulton, J. P., Shearer, S., & Ozkan, E. (2020). Remote sensing in agriculture—accomplishments, limitations, and opportunities. *Remote Sensing*, 12(22), 3783. <https://doi.org/10.3390/rs12223783>
- Kimes, D. S., Markham, B. L., Tucker, C. J., & McMurtrey III, J. E. (1981). Temporal relationships between spectral response and agronomic variables of a corn canopy. *Remote Sensing of Environment*, 11, 401-411. [https://doi.org/10.1016/0034-4257\(81\)90037-7](https://doi.org/10.1016/0034-4257(81)90037-7)
- Kobayashi, N., Tani, H., Wang, X., & Sonobe, R. (2020). Crop classification using spectral indices derived from Sentinel-2A imagery. *Journal of Information and Telecommunication*, 4(1), 67-90. <https://doi.org/10.1080/24751839.2019.1694765>
- Kosari, A., Sharifi, A., Ahmadi, A., & Khoshsima, M. (2020). Remote sensing satellite's attitude control system: rapid performance sizing for passive scan imaging mode. *Aircraft Engineering and Aerospace Technology*, 92(7), 1073-1083. <https://doi.org/10.1108/AEAT-02-2020-0030>
- Kumar, L., & Mutanga, O. (2018). Google Earth Engine applications since inception: Usage, trends, and potential. *Remote Sensing*, 10(10), 1509. <https://doi.org/10.3390/rs10101509>
- Lobell, D. B., Di Tommaso, S., You, C., Djima, I. Y., Burke, M., & Kilic, T. (2019). Sight for sorghums: Comparisons of satellite-and ground-based sorghum yield estimates in Mali. *Remote Sensing*, 12(1), 100. <https://doi.org/10.3390/rs12010100>
- Loukika, K. N., Keesara, V. R., & Sridhar, V. (2021). Analysis of land use and land cover using machine learning algorithms on google earth engine for Munneru River Basin, India. *Sustainability*, 13(24), 13758. <https://doi.org/10.3390/su132413758>
- Löw, F., Michel, U., Dech, S., & Conrad, C. (2013). Impact of feature selection on the accuracy and spatial uncertainty of per-field crop classification using support vector machines. *ISPRS Journal of Photogrammetry and Remote Sensing*, 85, 102-119. <https://doi.org/10.1016/j.isprsjprs.2013.08.007>
- Luo, C., Liu, H., Lu, L., Liu, Z., Kong, F., & Zhang, X. (2021). Monthly composites from Sentinel-1 and Sentinel-2 images for regional major crop mapping with Google Earth Engine. *Journal of Integrative Agriculture*, 20(7), 1944-1957. [https://doi.org/10.1016/S2095-3119\(20\)63329-9](https://doi.org/10.1016/S2095-3119(20)63329-9)
- Mahalanobis, A., Kumar, B. V., & Sims, S. R. F. (1996). Distance-classifier correlation filters for multiclass target recognition. *Applied Optics*, 35(17), 3127-3133. <https://doi.org/10.1364/AO.35.003127>
- Malenovsky, Z., Rott, H., Cihlar, J., Schaepman, M. E., García-Santos, G., Fernandes, R., & Berger, M. (2012). Sentinels for science: Potential of Sentinel-1, -2, and -3 missions for scientific observations of ocean, cryosphere, and land. *Remote Sensing of Environment*, 120, 91-101. <https://doi.org/10.1016/j.rse.2011.09.026>
- Marx, A. J., & Loboda, T. V. (2013). Landsat-based early warning system to detect the destruction of villages in Darfur, Sudan. *Remote Sensing of Environment*, 136, 126-134. <https://doi.org/10.1016/j.rse.2013.05.006>
- Mashala, M. J., Dube, T., Mudereri, B. T., Ayisi, K. K., & Ramudzuli, M. R. (2023). A Systematic Review on Advancements in Remote Sensing for Assessing and Monitoring Land Use and Land Cover Changes Impacts on Surface Water Resources in Semi-Arid Tropical Environments. *Remote Sensing*, 15(16), 3926. <https://doi.org/10.3390/rs15163926>
- Maxwell, A. E., Warner, T. A., & Fang, F. (2018). Implementation of machine-learning classification in remote sensing: An applied review. *International Journal of Remote Sensing*, 39(9), 2784-2817. <https://doi.org/10.1080/01431161.2018.1433343>
- Mazarire, T. T., Ratschiedana, P. E., Nyamugama, A., Adam, E., & Chirima, G. (2020). Exploring machine learning algorithms for mapping crop types in a heterogeneous agriculture landscape using Sentinel-2 data. A case study of Free State Province, South Africa. *South African Journal of Geomatics*, 9(2), 333-347. <https://doi.org/10.4314/sajg.v9i2.22>
- Mirzabaev, A., Sakketa, T., Sylla, M. B., Dimobe, K., Sanfo, S., Admassie, A., Abebaw, D., Coulibaly, O. N., Adamou, R., Ibrahim, B., Bonkaney, A. L., Seyni, A. A., Idrissa, M., Olayide, O. E., Faye, A., Dièye, M., Diakhaté, P. B., Beye, A., Sall, M., von Braun, J. (2021). Land, Climate, Energy, Agriculture and Development in the Sahel: Synthesis paper of case studies under the Sudano-Sahelian Initiative for Regional Development, Jobs, and Food Security. <https://doi.org/10.2139/ssrn.3769155>
- Mohammed, A., Zhang, K., Kabenge, M., Keesstra, S., Cerdà, A., Reuben, M., Elbashier, M. M. A., Dalson, T., & Ali, A. A. S. (2018). Analysis of drought and vulnerability in the North Darfur region of Sudan. *Land Degradation & Development*, 29(12), 4424-4438. <https://doi.org/10.1002/ldr.3180>
- Nitze, I., Schulthess, U., & Asche, H. (2012). Comparison of machine learning algorithms random forest, artificial neural network and support vector machine to maximum likelihood for supervised crop type classification. *Proceedings of the 4th GEOBIA, Rio de Janeiro, Brazil*, 79, 35-40.
- Osman, A. (2017). Climate Change and Drylands Farming in Sudan: Trend, Impact and Adaptation. *UK Journal of Natural Resources and Environmental Studies Special Issue Oct, 2017*, 1-10.
- Osman, A. M. K. (2012). *Agricultural change, land and violence: An examination of the region of Darfur, Sudan*. Doctoral Dissertation, Tufts University.
- Osman, A. M. K., & Cohen, M. J. (2014). *We No Longer Share the Land: Agricultural change, land, and violence in Darfur*. Oxford, UK: Oxfam International.
- Ouzemou, J.-E., El Harti, A., Lhissou, R., El Moujahid, A., Bouch, N., El Ouazzani, R., Bachaoui, E. M., & El Ghmari, A. (2018). Crop type mapping from pansharpened Landsat 8 NDVI data: A case of a highly fragmented and intensive agricultural system. *Remote Sensing Applications: Society and Environment*, 11, 94-103. <https://doi.org/10.1016/j.rsase.2018.05.002>
- Ozdogan, M. (2010). The spatial distribution of crop types from MODIS data: Temporal unmixing using Independent Component Analysis. *Remote Sensing of Environment*, 114(6), 1190-1204. <https://doi.org/10.1016/j.rse.2010.01.006>
- Paliwal, A., & Jain, M. (2020). The accuracy of self-reported crop yield estimates and their ability to train remote sensing algorithms.

- Frontiers in Sustainable Food Systems*, 4, 25. <https://doi.org/10.3389/fsufs.2020.00025>
- Pasternak, M., & Pawluszek-Filipiak, K. (2022). The evaluation of spectral vegetation indexes and redundancy reduction on the accuracy of crop type detection. *Applied Sciences*, 12(10), 5067. <https://doi.org/10.3390/app12105067>
- Pimple, U., Simonetti, D., Sitthi, A., Pungkul, S., Leadprathom, K., Skupek, H., Som-ard, J., Gond, V., & Towprayoon, S. (2018). Google earth engine based three decadal landsat imagery analysis for mapping of mangrove forests and its surroundings in the trat province of Thailand. *Journal of Computer and Communications*, 6(1), 247-264. <https://doi.org/10.4236/jcc.2018.61025>
- Praticò, S., Solano, F., Di Fazio, S., & Modica, G. (2021). Machine learning classification of mediterranean forest habitats in google earth engine based on seasonal sentinel-2 time-series and input image composition optimisation. *Remote Sensing*, 13(4), 586. <https://doi.org/10.3390/rs13040586>
- Qader, S. H., Dash, J., Alegana, V. A., Khwarahm, N. R., Tatem, A. J., & Atkinson, P. M. (2021). The role of earth observation in achieving sustainable agricultural production in arid and semi-arid regions of the world. *Remote Sensing*, 13(17), 3382. <https://doi.org/10.3390/rs13173382>
- Qader, S. H., Utazi, C. E., Priyatikanto, R., Najmaddin, P., Hama-Ali, E. O., Khwarahm, N. R., Tatem, A. J., & Dash, J. (2023). Exploring the use of Sentinel-2 datasets and environmental variables to model wheat crop yield in smallholder arid and semi-arid farming systems. *Science of The Total Environment*, 869, 161716. <https://doi.org/10.1016/j.scitotenv.2023.161716>
- Qian, Y., Zhou, W., Yan, J., Li, W., & Han, L. (2014). Comparing machine learning classifiers for object-based land cover classification using very high resolution imagery. *Remote Sensing*, 7(1), 153-168. <https://doi.org/10.3390/rs70100153>
- Rao, P., Zhou, W., Bhattarai, N., Srivastava, A. K., Singh, B., Poonia, S., Lobell, D. B., & Jain, M. (2021). Using Sentinel-1, Sentinel-2, and Planet imagery to map crop type of smallholder farms. *Remote Sensing*, 13(10), 1870. <https://doi.org/10.3390/rs13101870>
- Richardson, A. J., & Wiegand, C. L. (1977). Distinguishing vegetation from soil background information. *Photogrammetric Engineering and Remote Sensing*, 43, 1541-1552.
- Rodriguez-Galiano, V. F., Ghimire, B., Rogan, J., Chica-Olmo, M., & Rigol-Sanchez, J. P. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 67, 93-104. <https://doi.org/10.1016/j.isprsjprs.2011.11.002>
- Rouse, J. W., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. *NASA Special Publication*, 351, 309.
- Roy, D. P., & Dikshit, O. (1994). Investigation of image resampling effects upon the textural information content of a high spatial resolution remotely sensed image. *International Journal of Remote Sensing*, 15(5), 1123-1130. <https://doi.org/10.1080/01431169408954146>
- Rwanga, S. S., & Ndambuki, J. M. (2017). Accuracy assessment of land use/land cover classification using remote sensing and GIS. *International Journal of Geosciences*, 8(4), 611-622. <https://doi.org/10.4236/ijg.2017.84033>
- Sánchez-Espinosa, A., & Schröder, C. (2019). Land use and land cover mapping in wetlands one step closer to the ground: Sentinel-2 versus landsat 8. *Journal of Environmental Management*, 247, 484-498. <https://doi.org/10.1016/j.jenvman.2019.06.084>
- Sapkota, T. B., Jat, M. L., Jat, R. K., Kapoor, P., & Stirling, C. (2016). Yield estimation of food and non-food crops in smallholder production systems. In T. S. Rosenstock, M. C. Rufino, K. Butterbach-Bahl, L. Wollenberg & M. Richards (Eds.), *Methods for measuring greenhouse gas balances and evaluating mitigation options in smallholder agriculture* (pp. 163-174) Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-319-29794-1_8
- Segarra, J., Buchailot, M. L., Araus, J. L., & Kefauver, S. C. (2020). Remote sensing for precision agriculture: Sentinel-2 improved features and applications. *Agronomy*, 10(5), 641. <https://doi.org/10.3390/agronomy10050641>
- Shaharum, N. S. N., Shafri, H. Z. M., Ghani, W. A. W. A. K., Samsatli, S., Al-Habshi, M. M. A., & Yusuf, B. (2020). Oil palm mapping over Peninsular Malaysia using Google Earth Engine and machine learning algorithms. *Remote Sensing Applications: Society and Environment*, 17, 100287. <https://doi.org/10.1016/j.rsase.2020.100287>
- Shelestov, A., Lavreniuk, M., Kussul, N., Novikov, A., & Skakun, S. (2017). Exploring Google Earth Engine platform for big data processing: Classification of multi-temporal satellite imagery for crop mapping. *Frontiers in Earth Science*, 5, 17. <https://doi.org/10.3389/feart.2017.00017>
- Shetty, S. (2019). *Analysis of machine learning classifiers for LULC classification on Google Earth Engine*. Masters Thesis, University of Twente.
- Sinha, R., Quirós, J. J., Sankaran, S., & Khot, L. R. (2022). High resolution aerial photogrammetry based 3D mapping of fruit crop canopies for precision inputs management. *Information Processing in Agriculture*, 9(1), 11-23. <https://doi.org/10.1016/j.inpa.2021.01.006>
- Sonobe, R., Tani, H., Wang, X., Kobayashi, N., & Shimamura, H. (2014). Parameter tuning in the support vector machine and random forest and their performances in cross-and same-year crop classification using TerraSAR-X. *International Journal of Remote Sensing*, 35(23), 7898-7909. <https://doi.org/10.1080/01431161.2014.978038>
- Sothe, C., de Almeida, C. M., Liesenberg, V., & Schimalski, M. B. (2017). Evaluating Sentinel-2 and Landsat-8 data to map successional forest stages in a subtropical forest in Southern Brazil. *Remote Sensing*, 9(8), 838. <https://doi.org/10.3390/rs9080838>
- Sujud, L., Jaafar, H., Hassan, M. A. H., & Zurayk, R. (2021). Cannabis detection from optical and RADAR data fusion: A comparative analysis of the SMILE machine learning algorithms in Google Earth Engine. *Remote Sensing Applications: Society and Environment*, 24, 100639. <https://doi.org/10.1016/j.rsase.2021.100639>
- Thenkabail, P. S., Smith, R. B., & De Pauw, E. (2000). Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. *Remote Sensing of Environment*, 71(2), 158-182. [https://doi.org/10.1016/S0034-4257\(99\)00067-X](https://doi.org/10.1016/S0034-4257(99)00067-X)
- Tran, K. H., Zhang, H. K., McMaine, J. T., Zhang, X., & Luo, D. (2022). 10 m crop type mapping using Sentinel-2 reflectance and 30 m cropland data layer product. *International Journal of Applied Earth Observation and Geoinformation*, 107, 102692. <https://doi.org/10.1016/j.jag.2022.102692>
- Vuolo, F., Neuwirth, M., Immitzer, M., Atzberger, C., & Ng, W.-T. (2018). How much does multi-temporal Sentinel-2 data improve crop type classification? *International Journal of Applied Earth Observation and Geoinformation*, 72, 122-130. <https://doi.org/10.1016/j.jag.2018.06.007>
- Xie, G., & Niculescu, S. (2022). Mapping crop types using sentinel-2 data machine learning and monitoring crop phenology with sentinel-1 backscatter time series in pays de Brest, Brittany, France. *Remote Sensing*, 14(18), 4437. <https://doi.org/10.3390/rs14184437>
- Xue, J., & Su, B. (2017). Significant remote sensing vegetation indices: A review of developments and applications. *Journal of Sensors*, 2017(1), 353691. <https://doi.org/10.1155/2017/1353691>
- Zhang, H., Kang, J., Xu, X., & Zhang, L. (2020). Accessing the temporal and spectral features in crop type mapping using multi-temporal Sentinel-2 imagery: A case study of Yi'an County, Heilongjiang province, China. *Computers and Electronics in Agriculture*, 176, 105618. <https://doi.org/10.1016/j.compag.2020.105618>