Cross-Regional Transferability of AI Crop-Type Mapping: Insights and Challenges

Valinho António
vantonio@rsif-paset.org
African Centre of Excellence in Internet of Things, University of Rwanda
Kigali, Rwanda

Geoffrey Kimani
Carnegie Mellon University-Africa
Kigali, Rwanda
gnkimani@andrew.cmu.edu

Eric Umuhoza
Carnegie Mellon University-Africa
Kigali, Rwanda
eumuhoza@andrew.cmu.edu

Moise Busogi
Carnegie Mellon University-Africa
Kigali, Rwanda
mbusogi@andrew.cmu.edu

ABSTRACT

Accurate crop discrimination is vital for effective agricultural planning and sustainability management, especially in regions like Sub-Saharan Africa (SSA), where small-scale farming predominates and ground data is scarce. Conducting field surveys in SSA is challenging due to labor and cost constraints, as well as logistical and political barriers. This paper explores the feasibility of transferring crop-type classification models between regions with similar crops. Utilizing data from Baraouéli and Karamoja, collected from Source Cooperative, we trained multi-layer perceptron (MLP), random forest (RF), and support vector machine (SVM) classifiers using Sentinel-2 imagery. These models were then evaluated and applied to cross-crop type classification between Baraouéli (Mali) and Karamoja (Uganda) to assess the transferability of machine learning models. While the models demonstrated strong local performance, achieving high overall accuracy in their respective regions, their performance declined when transferred between regions. However, focusing solely on specific crops such as maize and sorghum improved the models' performance, albeit with reduced accuracy compared to local classifications. The study suggests that incorporating additional features such as texture, DEM, crop height, and weather data could enhance the adaptability of classifiers between regions. These findings highlight the potential for developing transferable models within SSA to address challenges related to limited ground surveyed data, providing valuable insights for researchers and policymakers.

CCS CONCEPTS

• **Computing methodologies** → *Machine learning approaches*; *Object identification*; • **Applied computing** → *Agriculture*; *Environmental sciences.*

KEYWORDS

Machine learning, crop types mapping, Sentinel-2 images, transfer learning.

ACM Reference Format:
Valinho António, Geoffrey Kimani, Eric Umuhoza, and Moise Busogi. 2024. Cross-Regional Transferability of AI Crop-Type Mapping: Insights and Challenges. In *International Conference on Information Technology for Social Good (GoodIT '24), September 04–06, 2024, Bremen, Germany*. ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3677525.3678696

GoodIT '24, September 04–06, 2024, Bremen, Germany

ACM ISBN 979-8-4007-1094-0/24/09
https://doi.org/10.1145/3677525.3678696

1 INTRODUCTION

The Food and Agriculture Organization (FAO) predicts that the demand for food production will increase by approximately 60%[11], primarily due to population growth [11], which is estimated to increase by around three billion people by 2050 [5]. This surge in demand is expected to triple in sub-Saharan Africa (SSA) over the next 30 years, requiring an additional 140 million hectares (ha) of land for agricultural use [31]. To decrease this pressure of agriculture expansion which is one of the drivers for forest loss, and, at the same time, to meet the escalating food demand in SSA, it is essential to increase the productivity of the most consumed cereals, such as maize and sorghum, particularly in regions heavily reliant on smallholder farming [5], using a sustainable agricultural practices. Accurate and timely crop discrimination plays a critical role in achieving these goals by enabling targeted agricultural interventions, effective agriculture planning and management, effective resource allocation, and precise decision-making [32].

Traditionally, most SSA governments track agriculture information, like crop types, acreage, and productivity, through agriculture surveys at the district level, which are mostly outdated [5]. Moreover, these surveys are difficult, time-consuming, and labor-intensive to conduct. Furthermore, this method may lack spatial coverage, making it difficult to capture the variability of crops across large agricultural landscapes. On the other hand, high-resolution satellite imagery presents a transformative opportunity: they can provide a comprehensive and synoptic view of agricultural landscapes, capturing a massive amount of data related to spectral information that reflects diverse crop characteristics. The ability to leverage this vast amount of data lies in the development of robust and efficient machine learning (ML) algorithms that excel in processing and interpreting the data, allowing rapid and precise crop type discrimination [35].

The integration of machine learning and satellite images plays a pivotal role in discerning subtle differences between crop types and it enables analyzing large agricultural areas efficiently. So, this timely and reliable information about crop types and their spatial distribution can help decision-makers in optimizing resource allocation, such as agricultural extensionists, fertilizers, and pesticides, to the farmers according to their needs and to estimate crop production [32]. However, even with the imperative of precise crop identification in SSA, due to its profound impact on food security and agricultural sustainability, there are few studies done in this context where the majority of the population heavily relies on small farms (< 2 *ha*) and ground data is almost absent.

The ground survey data which is scarce in many SSA countries is paramount for crop type classification using supervised learning, but field surveys are labor intensive and very expensive, and in some parts of SSA are impossible to conduct due to political or logistical reasons[30]. Consequently, it can be difficult to map crop types with high accuracy in this region[12]. Nevertheless, many attempts have been made to collect samples from publicly available and annual updated existing country-wide operational crop types mapping, such as Crop Data Layer (CDL) from the US Department of Agriculture (USDA) and Agriculture and Agri-Food Canada's Annual Crop Inventory (AAFC), to train a classifier and then applying it to other site and or year. For example, Zhong et al.[38] improved the time for early season crop type discrimination using reference data of previous years from CDL and they achieved satisfactory results. Di Tommaso et al. have shown that the synergy between NASA's Global Ecosystem Dynamics Investigation (GEDI) spaceborne LiDAR and Sentinel-2 (GEDI-S2) is very promising for crop types transfer model, as they found that GEDI-S2 performed nearly well like the models trained on local reference samples, with accuracies between 87% and 90% throughout US, France, and China. The authors used CDL, Registre Parcellaire Graphique (RPG) 2019 dataset for France[8], and 2019 crop type produced by You et al.[34] as reference datasets.

However, to the best of our knowledge, there is no product like CDL and AAFC in SSA region and, as in the local crop types mapping, these attempts are more concentrated in developed countries such as China, Canada, and USA, which hamper the decision-making related to food security. So, starting from the principle that the phenological and growing patterns of the same crop are more similar than that of different crops, even between different regions of the world [12]. Based on this assumption, we can hypothesize that an ML model trained in one site and year can be applied successfully to another place in the same or different year, ignoring the rapid inter-annual seasonal changes and the differences in regional climates which can affect the crop progress, thereby the spectral response[17, 30].

Our objective is to leverage reference samples from Baraouéli and Karamoja, accessible via Source Cooperative (formerly Radiant MLHub), to train three widely utilized machine learning models for crop type discrimination using Sentinel-2 imagery: multi-layer perceptron (MLP), random forest (RF), and support vector machine (SVM) in a single location (either Baraouéli or Karamoja). Subsequently, we will assess and deploy these models in a different location and year from the training data to investigate the feasibility of transferring a machine learning model trained in Mali to accurately identify crop types in Uganda, and vice versa, relying solely on spectral information.

2 MATERIALS AND METHODS

2.1 Study regions

This study considers two regions—Baraouéli, Mali, and Karamoja, Uganda—their localization and topography are depicted in Figure 1. These two regions are chosen first because of the availability of reference data and secondly due to their similarities in climate which leads to almost the same growing season for most of the crops such as maize and sorghum [21].

Climate. Baraouéli is characterized by a semi-arid climate with two distinct seasons: a rainy season from June to October and a dry season from November to May. The average temperature of the winter is 25.6 °C and the summer average temperature is above 30°C. Annual precipitation varies between 600 mm and 900 mm, ranging from 0 mm in the driest months (February—December) to 270 mm in the wettest (August). Similarly, Karamoja experiences a semi-arid with distinct wet and dry seasons [7]. The dry season is a prolonged one, typically lasting from December to March and the Wet season usually occurs from April to November, with the heaviest rainfall falling between June and September. Temperatures averaging 30.5 °C (maximum) and 16.5 °C (minimum). The annual precipitation varies between 500 mm to 1000 mm and its distribution is uneven with some areas receiving more precipitation than others [7].

Agriculture profile. Like in many SSA countries, agriculture is dominated by smallholder farms (fields between 0.5 and 5 ha) in both regions, and sorghum, maize, and finger millet are the most dominant crops [7, 9].

Topography. The Baraouéli's terrain is mostly flat with some undulating areas and it has an average elevation of 316.08 m above the sea [10], while Karamoja is situated on a large plateau with hills and plain terrains, and with an average elevation of approximately 1287.05 m above the sea level (Figure 1).

2.2 Data collection and preparation

This study uses a combination of two spaceborne sensors: Sentinel-2 and PlanetScope. Multiple dates of images from Sentinel-2 Level 1C Top-Of-Atmosphere data (TOA) with less than 20% of cloud cover were downloaded from May to September (grown season of maize and sorghum) of 2017 (Karamoja) and 2019 (Baraouéli) using Google Earth Engine (GEE). The bands with 20 m spatial resolution (B5, B6, B8A, B11, and B12) were resampled to 10 m to match the bands B2, B3, B4, and B8 (Table 1) and the TOA images were cloud masked using built-in GEE algorithm with a Quality Assessment-60 (QA60) band that controls clouds and cirrus pixels, and visually observation of the masked images. Then, we chose one image for each month in Baraouéli because it covered the entire region without severe cloud problems and we ended up having four images (Table 1) for the entire grown season. Whereas, for Karamoja, we did a mosaic of all imagery available for each month to cover the entire region, thereby we retrieved 18 images from May, 47 from July,

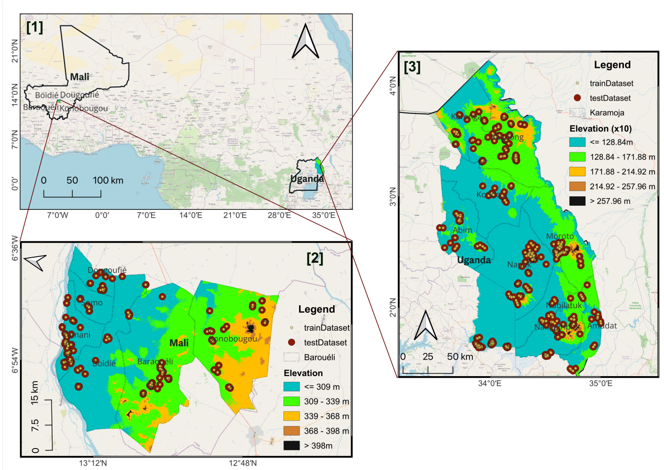

Figure 1: Study regions: (1) Baraouéli located in Mali and Karamoja located in Uganda, (2) Extent of Baraouéli commune with its topography and in-situ data, (3) Extent of Karamoja with its topography and in-situ data

56 from August, and 42 images from September (Table 1). Images from June were not included due to their unavailability in Baraouéli. However, it is worth noting that we didn't convert the images to Surface-Reflactance (SR) due to several challenges including the lack of an efficient GEE algorithm for radiometric corrections, unavailability of the processor Sen2Cor for Sentinel-2 Level 2A (SR), and limited SR data in GEE for the study regions. In addition to the Sentinel-2 bands, we derived the six most used radiometric indices for crop type discrimination in SSA based on a systematic literature review we did: Meris Terrestrial Chlorophyll Index (MTCI), Normalized Vegetation Index (NDVI), Green Chlorophyll Vegetation Index (GCVI), Normalized Difference Water Index (NDWI), Soil Adjusted Vegetation Index (SAVI), and Enhanced Vegetation Index (EVI). The equations of these indices are reported in Table 1.

From PlanetScope Basemaps [23], we retrieved PS Tropical Normalized Analytic Biannual Archive quads from through Norway's International Climate and Forest Initiative (NICFI) [22] program in QGIS using Planet QGIS Plugin, then we composed a mosaic of the quads. Later, the mosaic was used to add polygons for four land use categories: tree, built, bare soil, and water which were not present in ground truth data for the two regions. During the adding process, we explored the false color composite and NDVI to extract these categories.

Finally, the reference datasets for the two study regions were collected from Source Cooperative [6], former Radiant MLHub. The Baraouéli Cercle ground truth data were 2019 Mali Crop-type Training Data for Machine Learning. This dataset was produced by the NASA Harvest team through field visits using Mobile Phone GPS with continuous point capture of the entire field boundary. The observed and recorded data in the ODK application were: no data, millet, maize, sorghum, and rice. Based on this ground survey, the labels identified were further vectorized over the Sentinel-2 grid and provided as raster files [20]. From the raster files, we joined them, converted them into shapefiles, and then extracted maize and sorghum as the target crops for our study because these are the labels available in the Karamoja in-situ data. We found 16 polygons for maize and 22 for sorghum. Moreover, we extracted training and testing sample points for each class and we reprojected the results from EPSG:4326-WGS 84 to EPSG:32629-WGS 84/ UTM zone 29N to match the Sentinel-2 data exported from GEE. We used random points in the polygons processing toolbox with 40 points per polygon and a minimum distance between the points of 2.5 meters, thereby we got 640 pixels for maize, 842 for sorghum, 774 points for trees, 887 for built-up areas, 720 for bare soil, and 786 pixels for water. From each of these points, we divided into 70% training and 30% testing (Figure 1) using a random selection algorithm. All these data preparation were done in QGIS 3.34.4-Prizren [25].

The Karamoja subregion reference data were Dalberg Data Insights Uganda Crop Classification collected by the Dalberg Data Insights team over a ground survey at the end of September 2017 using tablet-enabled GPS to capture a sample of a field in a quadrant way [15]. The recorded crop types were maize and sorghum. A data quality control was done by Radiant Earth Foundation to check the polygons using Sentinel-2 and Google basemap images of the crops growing season before their publication in Radiant MLHub (now Source Cooperative), so there were several polygons removed which overlapped with infrastructure or built-up areas [15]. The final data were vectorized and provided in geojson format. From the geojson files, we joined them and converted them into shapefiles, besides the separation between maize and sorghum to generate training and testing pixels easily. We found 108 polygons for maize and 126 for sorghum and their shapefiles were in the same projection as Sentinel-2 imagery, consequently, we did not do any reprojection. As in the Baraouéli dataset, we used random points in polygons and random selection algorithms to get pixels for machine learning model training and testing. As a result, we got 1620 pixels for maize, 1512 for sorghum, 1431 for trees, 1465 points for built-up areas, 1500 for bare soil, and 1318 pixels for water. These points were further divided into training (70%) and testing (30%) like in the Baraouéli data (Figure 1).

2.3 Data analysis

Supervised classification models in remote sensing have been extensively utilized to identify crop types globally [27, 28]. This study employs three commonly used techniques—random forests (RF), support vector machine (SVM), and deep learning (DL) with dense layer architecture—to classify crop types in SSA. The classifiers were applied to distinguish crops and other land use categories in Baraouéli (2019 growing season) and Karamoja (2017 growing season) using spectral bands and radiometric indices as inputs.

The RF is an ensemble technique that comprehends a set of decision trees created by drawing a subset of training data using a bootstrap aggregation and the final decision is obtained by majority votes of the decision trees and it has been successfully used in crop discrimination[18, 29]. Therefore, we employed the default parameters of RandomForestClassifier method from Python's scikit-learn package except for the number of decision trees (n_estimators)

Table 1: Sentinel-2 acquisition dates, bands, and spectral indices.

Region	Image date and quantity	Features	
		Bands and spatial resolution	Indice & equation
Baraouéli (Mali)	26/05/2019; 25/07/2019; 14/08/2019; 03/09/2019 (4 images).	B2, B3, B4, B8 (10 m); B5, B6, B8A, B11, B12 (20 m).	$MTCI = \frac{B8(NIR) - B5(RE1)}{B5 - B4(Red)}$
			$NDVI = \frac{B8 - B4}{B8 + B4}$
			$GCVI = \frac{B8}{B3(Green) - 1}$
Karamoja (Uganda)	02-25/05/2017 (18 images); 01-29/07/2017 (47 images); 03-28/08/2017 (56 images); 07-27/09/2017 (42 images).		$NDWI = \frac{B3 - B8}{B3 + B8}$
			$SAVI = \frac{B8 - B4}{(B8 + B4 + L) \times 1 + L}$
			$EVI = G \times \frac{B8 - B4}{(B8 + C1 \times B4) - (C2 \times B2 + 1)}$

which was set to 340 with the optimal accuracy after testing different numbers of decision trees from 100 to 500. The datasets were randomly split following a classific approach 70%-30% into a training and a test set, respectively. Moreover, we used out-of-bag (OOB)[4] accuracy to measure the RF model's overall accuracy.

An SVM is an ML algorithm that aims to find the optimal hyperplane that maximizes the distance between the closest training sample and the separating hyperplane in the feature space using different kernel functions[19]. It has been successfully used in crop identification in the context of SSA [2, 3]. In this study, we used the most common kernel function among the remote sensing community which is radial basis function (RBF) [2] within support vector classifier function (SVC) from Python's scikit-learn SVM with default parameters, excepting the cost which was set to be 50.

The DL is a subfield of ML that involves the use of neural networks with multiple processing layers to learn complex representations of data with multiple levels of abstraction [16]. There are different architectures of DL proposed for solving different problems such as multi-layer perceptron (MLP) and convolutional neural networks (CNN) which have achieved remarkable success in many domains including crop types discrimination[24, 33]. In this way, we developed and used an MLP for crop identification as our data were structured using Python's TensorFlow library with Dense and Sequential functions with Leaky ReLu[1] and Linear activations(Figure 2). The parameters were trained using SparseCategoricalCrossentropy Kera's loss function and Adam algorithm optimizer with integrated softmax output function to handle the numerical round-off error[1]. The learning rate for the optimizer was set to 0.01 for Baraouéli and 0.00001 for Karamoja after many parameter tuning attempts.

The efficacy of the models was assessed through two metrics: overall accuracy (OA) to determine how many pixels a model predicted well and f1-score, which measures precision and recall at once, to determine how well and completely a classifier does a prediction [26].

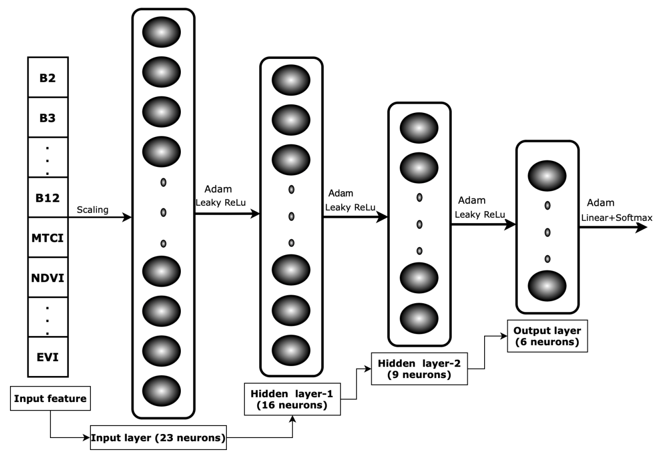

Figure 2: Multi-layer perceptron architecture.

3 RESULTS AND DISCUSSION

The results are presented by observing three experiments: results on local data, transferability from Baraouéli to Karamoja, and transferability from Karamoja to Baraouéli region.

3.1 Experiment 1: Performance of the classifiers on local data

Experiment 1 utilized multi-temporal spectral bands and radiometric indices throughout the entire growing season of maize and sorghum in both regions. Table 2 presents the classification results from Baraouéli, Mali. All models achieved high overall classification accuracies, with the deep learning model recording the lowest at 97%. The classes *Bare Soil* and *Built* exhibited excellent performance, achieving an F1-score of 1.00 in the deep learning model, and the random forest and support vector machine models. Karamoja's results are presented in Table 3. For this region, RF and SVM classifiers achieved high accuracies of 94% and 93%, respectively, whereas the

Table 2: Multi-temporal spectral bands and indices overall accuracy and f1-score for maize, sorghum and additional classes for all classifiers in Baraouéli (Mali)

Class	MLP		RF		SVM	
	OA	f1-Score	OA	f1-Score	OA	f1-Score
Maize	0.97	0.94	0.99	0.97	0.98	0.96
Sorghum		0.96		0.97		0.97
Tree		0.98		0.99		1.00
Built		0.99		1.00		1.00
Bare_Soil		1.00		1.00		0.98
Water		0.97		0.99		0.99

MLP model showed a lower accuracy of 82%. Importantly, while these models are the same as those used for Baraouéli, they show reduced accuracies with this dataset. The MLP model particularly struggled with the sorghum, tree, and built classes.

Table 3: Multi-temporal spectral bands and indices overall accuracy and f1-score for maize, sorghum and additional classes for all classifiers in Karamoja.

Class	MLP		RF		SVM	
	OA	f1-Score	OA	f1-Score	OA	f1-Score
Maize	0.82	0.82	0.94	0.95	0.93	0.94
Sorghum		0.78		0.93		0.92
Tree		0.79		0.93		0.92
Built		0.76		0.92		0.87
Bare_Soil		0.82		0.93		0.92
Water		0.98		1.00		0.99

Overall, the models performed very well, except MLP on Karamoja data, thereby they can be used in crop identification problems using spectral bands and radiometric indices.

The main goal of this study is to evaluate the transferability of these supervised models between different geographic regions and years to tackle the scarcity of in-situ data. To this end, in the experiments 3.2 and 3.3, the classifiers trained on data from Baraouéli are applied to Karamoja and vice versa.

3.2 Experiment 2: Transferability from Mali to Uganda

In this experiment, the models that were used for the classification of data in Baraouéli, Mali are applied to the Karamoja, Uganda dataset to assess their performance.

We observed that all the classifiers trained on the Mali dataset performed poorly (see Table 4) when applied to the Uganda dataset, revealing a lot of class pixels confusion between these two regions as shown in the confusion matrices in Figure 3. This can be related to the differences in Sentinel-2 backscatter values in each class due to climate and soil differences. For example, from the confusion matrices in Figure 3, we perceive that the MLP model with OA of 11% predicted almost all Ugandan pixels as water, the RF with OA of 11% and highest f1-score in sorghum of 26% mostly predicted the Ugandan pixels as sorghum and water, and that the SVM (17%) estimated all class pixels as bare soil.

Table 4: Overall accuracy and f1-score of transferred models from Mali to Uganda

Class	MLP		RF		SVM	
	OA	f1-Score	OA	f1-Score	OA	f1-Score
Maize	0.11	0.00	0.11	0.01	0.17	0.00
Sorghum		0.01		0.26		0.00
Tree		0.00		0.00		0.00
Built		0.00		0.00		0.00
Bare_Soil		0.00		0.00		0.27
Water		0.00		0.05		0.00

The detailed examination of class-specific classification performance (f1-scores), as presented in Table 4, shows that both the MLP and SVM models underperformed compared to the RF model, which identified sorghum with an f1-score of 26%. Given the notable performance of RF in predicting sorghum, we conducted a binary classification focusing solely on crop types to further analyze the

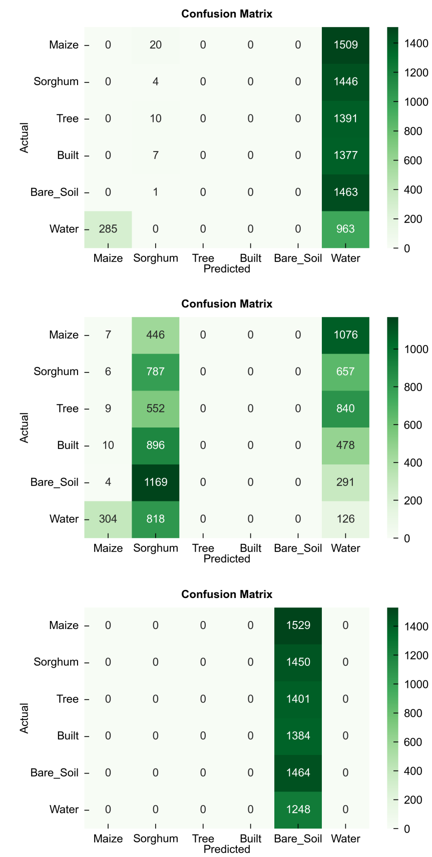

Figure 3: Confusion matrices of MLP (left), RF (middle), SVM (right) trained with Mali dataset and tested with Uganda data

behavior of the models. As presented in Table 5, MLP demonstrated the highest OA of 53%, which is reasonably better when compared to both RF and SVM, which each registered an OA of 49%.

It is important to note that sorghum was detected with an F1-score above 64%, indicating that all models have potential for effective use. Further, integrating additional features such as texture (from GLCM), climatic data (temperature and precipitation), alongside Digital Elevation Models (DEM) and soil type information, could significantly enhance the accuracy and reliability of these classifiers in identifying sorghum.

Table 5: Overall accuracy and f1-score of transferred models from Mali to Uganda using crops only

Class	MLP		RF		SVM	
	OA	f1-Score	OA	f1-Score	OA	f1-Score
Maize	0.53	0.23	0.49	0.00	0.49	0.00
Sorghum		0.67		0.65		0.65

From the confusion matrices (Figure 4), it's clear that all classifiers confused all maize pixels as sorghum, with the expectation of MLP which predicts 203 pixels as true class maize. This tells us the Sentinel-2 backscatter collected in Baraouéli for maize is very different from those collected in Karamoja. Whereas, for sorghum the spectral responses are very similar. As such, there is a need for further research in feature engineering to identify the features that will help to improve the discrimination of the crops, mainly the maize.

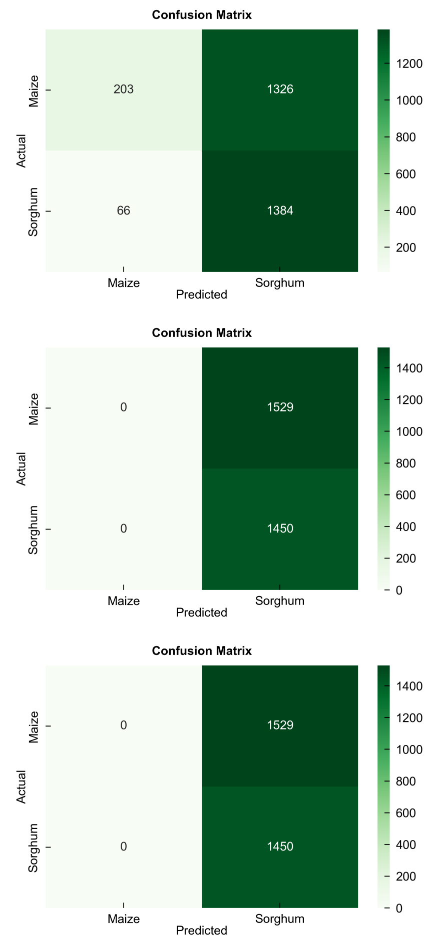

Figure 4: Confusion matrices from left to right:MLP, RF, and SVM trained with Mali dataset and tested on Uganda data using crops only.

3.3 Experiment 3: Transferability from Uganda to Mali

In this experiment, the models trained on Karamoja data are used to classify Baraouéli data. Similarly to the experiment in 3.2, the three classifiers also had poor performance: the deep learning model had an OA of 14%, 16% for RF, and SVM had an OA of 19% (Table 6).

These classification accuracy results are lower than those of the local models, indicating the differences in the spectral responses of the classes between Karamoja and Baraouéli. Nevertheless, different from experiment 2, between the crops the RF had zero predicted pixels as sorghum, same as SVM, but the highest f1-score was 28% achieved by MLP for sorghum, indicating that few sorghum feature values in Karamoja are similar to those in Baraouéli.

Table 6: Overall accuracy and f1-score of transferred models from Uganda to Mali

Class	MLP		RF		SVM	
	OA	f1-Score	OA	f1-Score	OA	f1-Score
Maize	0.14	0.00	0.16	0.00	0.19	0.00
Sorghum		0.28		0.00		0.00
Tree		0.13		0.00		0.00
Built		0.20		0.05		0.32
Bare_Soil		0.10		0.27		0.00
Water		0.00		0.00		0.00

The class spectral values differences were highlighted by Figure 5 with pixels distribution structure different from those predicted using model trained in Baraouéli with Karamoja dataset, as here the SVM estimated all pixels as built instead of bare soil, the RF considered almost all spectral responses as bare soil to detriment of water and sorghum, and finally the MLP had no clear concentration of the pixels, but it worth noting that there is no pixel predict as maize. One plausible reason for these misclassifications is that the spectral differences are bigger between the two regions, thereby the features are less informative.

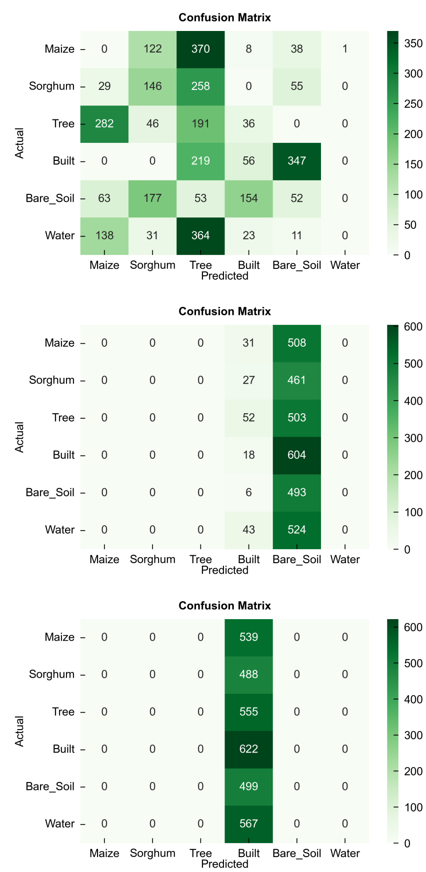

Figure 5: Confusion matrices of MLP (left), RF (middle), SVM (right) trained with Uganda dataset and tested with Mali data.

Similar to experiment two, we have applied a binary classification using only maize and sorghum to assess the performance of the models with crops only. Unlike in the former experiment, the MLP had the lowest OA of 42%, followed by RF with 43%, and the highest OA of 48% was achieved by SVM (Table 7). However, it is worth highlighting that the MLP had a reasonable identification of maize with an f1-score of 40% which is the highest score in all experiments for this crop (Table 7). Additionally, sorghum, as in the previous experiment, was estimated with the highest f1-score of 64% by SVM, followed by RF with 59%, and MLP had the lowest f1-score of 44% (Table 7). These results, similar to those in experiment two, indicate limited similarity in spectral response or phenological patterns due to varying weather conditions and crop management practices across the two regions and years [12, 14].

Table 7: Overall accuracy and f1-score of transferred models from Uganda to Mali using crops only

Class	MLP		RF		SVM	
	OA	f1-Score	OA	f1-Score	OA	f1-Score
Maize	0.42	0.40	0.43	0.05	0.48	0.00
Sorghum		0.44		0.59		0.64

To identify where the greatest confusion between class pixels occurred, we analyzed the confusion matrices for each classifier, as illustrated in Figure 6. We observed that similar to the findings in experiment two (see Figure 4), RF and SVM classifiers predominantly misclassified pixels as sorghum. Conversely, the MLP model exhibited a broader distribution of spectral responses among the crops, with substantial confusion between the spectral responses of sorghum and maize, and vice versa. Additionally, the MLP model frequently misidentified maize spectral responses as those of sorghum.

In summary, the three models (MLP, RF, and SVM) demonstrated strong local performance, achieving an average OA of 98% in Baraouéli (Mali) and 90% in Northern Uganda. However, when expanding the classification to include all classes, the transferability of the classifiers, as measured by OA and f1-score, diminished due to differences in the spectral responses. The differences are well depicted by Figure 7 which shows the spectral indices correlations of the two places on the top diagonal, highlighted by the orange line.

Notably, performance improvements were observed when the classification was limited to just two crops: maize and sorghum. This enhancement was primarily due to the consistency in the spectral responses of sorghum across both regions and different years. It is well established that if the crop conditions in the reference and test sites are similar, then the transfer of reference knowledge can yield high accuracy, and the reverse is also true[13]. Consequently, the MLP and RF models could potentially be adapted from Karamoja to Baraouéli and vice versa to specifically identify sorghum, albeit with a slight drop in accuracy compared to local classifications. While incorporating additional features such as texture, DEM, crop height (derived from Lidar data)[27], temperature, and precipitation could further enhance the performance of these transferred models, these are beyond the current study's scope. Moreover, as Lin et al.[17] suggest, integrating data from other sensors like Sentinel-1, PlanetScope, and SkySat could capture additional features that enhance the classification accuracy.

Similar findings were reported by Wang et al.[29], who observed that the similarity in Growing Degree Days (GDD) between two

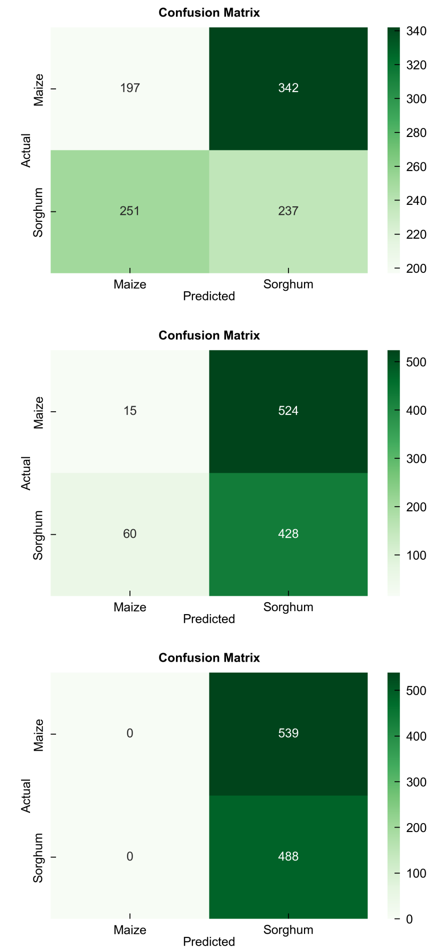

Figure 6: Confusion matrices from left to right: MLP, RF, and SVM trained with Uganda dataset and tested on Mali data using crops only.

regions correlates positively with the performance of machine learning models, particularly Random Forest. However, they noted that GDD may be less applicable in regions where temperature is not a critical factor for crop growth. Thus, it is recommended to explore alternative metrics such as precipitation, Digital Elevation Models (DEM), and crop height. These factors could prove invaluable in Sub-Saharan Africa, where in-situ data are scarce, making it challenging to train supervised models. Further, Zhang et al. [36]—in their study that used CNN, RF, and XGBoost algorithms alongside Sentinel-2 imagery to distinguish between rice and non-rice fields— discovered that these classifiers when transferred from one region to another exhibited lower accuracy than those calibrated and validated locally. Furthermore, they recognized that the generalizability of machine learning models remains a challenge because even slight variations in data patterns can lead to misclassifications: if the testing dataset deviates from the training samples, the model's performance will degrade[37]. This issue underpins the observed decline in classification accuracy when transferring classifiers across different regions

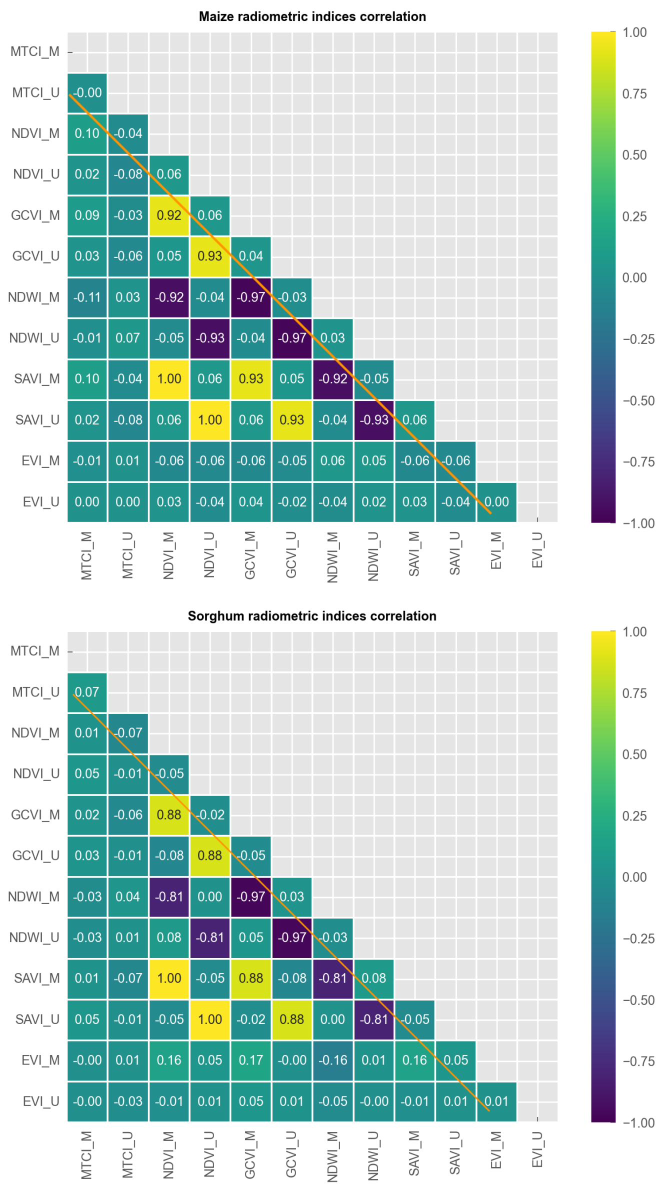

Figure 7: Radiometric indices correlation between Baraouéli and Karamoja for August: maize on the left and sorghum on the right.

and years. This topic is currently a vibrant subject of debate within the remote sensing community. Some researchers, holding a positivist view, advocate that a model trained in one region can be applicable elsewhere and in different years[27, 37], while others assert the contrary, pointing out that variations in crop spectral responses driven by environmental differences such as climate and soil type prevent such applicability [30].

4 CONCLUSIONS

This paper evaluated the transferability of MLP (DL), RF, and SVM classifiers for crop types discrimination between Baraouéli (Mali) and Karamoja (Northern Uganda) in different years using only spectral bands and radiometric indices. Training dataset were retried on Source Cooperative for the two regions. Therefore, the followings are the key conclusions: (i) the three models (MLP, RF, and SVM) demonstrated strong local performance, achieving an average OA of 98% in Baraouéli and 90% in Karamoja. However, the accuracies of the transferred classifiers when using all catergories, as measured by OA and f1-score, were diminished drastically; (ii) using only the crops (maize and sorghum), the performance of the models were improved partially due to the consistency in the spectral responses of maize and sorghum across the two regions in different years; (iii) the MLP and RF models could potentially be adapted from Northern Uganda to Baraouéli and vice versa to specifically identify sorghum, albeit with a slight drop in accuracy compared to local classifications; (iv) incorporating additional features such as texture, DEM, crop height, temperature, and precipitation could further enhance the performance of these transferred models.

ACKNOWLEDGMENTS

This work is part of a PhD thesis supported by the PASET-Regional Scholarship and Innovation Funds and hosted at the African Centre of Excellence in Internet of Things, University of Rwanda. The work was also partially supported by the National Council of Science and Technology (NCST), under the grant No. NCST-NRIF/ RIC-R&D-PHASE I/08 /05/2022. We also thank Source Cooperative for providing the Baraouéli (Mali) and Northern Uganda data and NICFI for providing PlanetScope basemaps.

REFERENCES

[1] Martín Abadi et al. 2016. TensorFlow: A System for Large-Scale Machine Learning. In *12th USENIX symposium on operating systems design and implementation (OSDI 16)*. USENIX Association, 265–283.

[2] Ghali Abubakar et al. 2020. Mapping maize fields by using multi-temporal Sentinel-1A and Sentinel-2A images in Makarfi, Northern Nigeria, Africa. *Sustainability* 12, 6 (2020), 2539. https://doi.org/10.3390/su12062539

[3] Tunrayo R Alabi, Julius Adewopo, Ojo Patrick Duke, and P Lava Kumar. 2022. Banana mapping in heterogenous smallholder farming systems using high-resolution remote sensing imagery and machine learning models with implications for banana bunchy top disease surveillance. *Remote Sensing* 14, 20 (2022), 5206. https://doi.org/10.3390/rs14205206

[4] L. Breiman. 2001. Random forests. *Machine learning* 45 (2001), 5–32. https://doi.org/10.1023/A:1010933404324

[5] Marshall Burke and David B. Lobell. 2017. Satellite-based assessment of yield variation and its determinants in smallholder African systems. *Proceedings of the National Academy of Sciences* 114, 9 (2017), 2189–2194. https://doi.org/10.1073/pnas.1616919114

[6] Source Cooperative. n.d. Source Cooperative. Source Cooperative, https://source.coop. (Accessed: 03 May 2024).

[7] A. Cullis. n.d. *Agricultural Development in Karamoja, Uganda: Recent Trends in Livestock and Crop Systems, and Resilience Impacts.* (Accessed: 02 May 2024).

[8] République Française (data.gouv.fr). n.d. *RPG 2019 - Registre Parcellaire Graphique.* (Accessed: 23 May 2024).

[9] Institut National de La Statistique. 2021. Frapport sur les principaux resultats definitifs du module agriculture intégré dans la phase cartographie du recensement général de la population et l'habitat, édition 5. Retrieved April 30, 2024 from https://www.instat-mali.org/storage/135/rapport-module-rga-rgph5_rap.pdf

[10] Conseils et Appui pour l'Education à la Base (CAEB). 2017. *Systeme de suivi-evaluation du fonds climat mali rapport d'activité annuelle.* Retrieved May 15, 2024 from https://mptf.undp.org/sites/default/files/documents/30000/rtpa_2017_rapport_activite_annuel_pam-caeb-sigi_1.pdf

[11] Food and Agriculture Organization (FAO). 2015. *Climate change and food security: risks and responses.* Retrieved May 13, 2024 from https://openknowledge.fao.org/

server/api/core/bitstreams/a4fd8ac5-4582-4a66-91b0-55abf642a400/content
[12] Pengyu Hao, Liping Di, Chen Zhang, and Liying Guo. 2020. Transfer Learning for Crop classification with Cropland Data Layer data (CDL) as training samples. *Science of The Total Environment* 733 (2020), 138869. https://doi.org/10.1016/j.scitotenv.2020.138869
[13] Pengyu Hao, Li Wang, Yulin Zhan, and Zheng Niu. 2016. Using moderate-resolution temporal NDVI profiles for high-resolution crop mapping in years of absent ground reference data: a case study of bole and manas counties in Xinjiang, China. *ISPRS International Journal of Geo-Information* 5, 5 (2016), 67. https://doi.org/10.3390/ijgi5050067
[14] Pengyu Hao, Li Wang, Yulin Zhan, Changyao Wang, Zheng Niu, and Mingquan Wu. 2016. Crop classification using crop knowledge of the previous-year: Case study in Southwest Kansas, USA. *European Journal of Remote Sensing* 49, 1 (2016), 1061–1077. https://doi.org/10.5721/EuJRS20164954
[15] Dalberg Data Insights. 2017. Dalberg Data Insights Uganda Crop Classification. DDI, September 2017 https://beta.source.coop/radiantearth/african-crops-uganda-01/. (Accessed: 03 May 2024).
[16] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. 2015. Deep learning. *nature* 521, 7553 (2015), 436–444. https://doi.org/10.1038/nature14539
[17] Chenxi Lin et al. 2022. Early-and in-season crop type mapping without current-year ground truth: Generating labels from historical information via a topology-based approach. *Remote Sensing of Environment* 274 (2022), 112994. https://doi.org/10.1016/j.rse.2022.112994
[18] Mmamokoma Grace Maponya, Adriaan Van Niekerk, and Zama Eric Mashimbye. 2020. Pre-harvest classification of crop types using a Sentinel-2 time-series and machine learning. *Computers and electronics in agriculture* 169 (2020), 105164. https://doi.org/10.1016/j.compag.2019.105164
[19] Farid Melgani and Lorenzo Bruzzone. 2004. Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transactions on geoscience and remote sensing* 42, 8 (2004), 1778–1790. https://doi.org/10.1109/TGRS.2004.831865
[20] C. L. Nakalembe, H. Ouedraogo, N. Diarra, and B. Kuzimbu. 2021. 2019 Mali Crop Type Training Data for Machine Learning. NASA-Harvest, 2021 https://beta.source.coop/umd/mali-crop-type/. (Accessed: 03 May 2024).
[21] U.S. Department of Agriculture (USDA). 2024. Foreign Agriculture Service: Crop Calendar Charts. Retrieved May 15, 2024 from https://ipad.fas.usda.gov/ogamaps/cropcalendar.aspx
[22] Planet. 2021. NICFI Data Program: user guide. Planet, June 2021 https://assets.planet.com/docs/NICFI_UserGuidesFAQ.pdf. (Accessed: 02 May 2024).
[23] Planet. 2024. PlanetScope. Planet, 02 May 2024 https://developers.planet.com/docs/data/planetscope/. (Accessed: 03 May 2024).
[24] Adriaan Jacobus Prins and Adriaan Van Niekerk. 2021. Crop type mapping using LiDAR, Sentinel-2 and aerial imagery with machine learning algorithms. *Geo-Spatial Information Science* 24, 2 (2021), 215–227. https://doi.org/10.1080/10095020.2020.1782776
[25] QGIS. 2023. QGIS User Guide. QGIS, 2023 https://docs.qgis.org/3.34/en/docs/user_manual/index.html. (Accessed: 03 May 2024).
[26] John A. Richards and Xiuping Jia. 2022. *Remote sensing digital image analysis: an introduction* (5 ed.). Springer.
[27] Stefania Di Tommaso, Sherrie Wang, Vivek Vajipey, Noel Gorelick, Rob Strey, and David B Lobell. 2023. Annual field-scale maps of tall and short crops at the global scale using gedi and sentinel-2. *Remote Sensing* 15, 17 (2023), 4123. https://doi.org/10.3390/rs15174123
[28] Gabriel Tseng, Ivan Zvonkov, Catherine Lilian Nakalembe, and Hannah Kerner. 2021. CropHarvest: A global dataset for crop-type classification. In *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2)*. https://openreview.net/forum?id=JtjzUXPEaCu
[29] Sherrie Wang, George Azzari, and David B Lobell. 2019. Crop type mapping without field-level labels: Random forest transfer and unsupervised clustering techniques. *Remote sensing of environment* 222 (2019), 303–317. https://doi.org/10.1016/j.rse.2018.12.026
[30] Yanan Wen et al. 2022. Mapping corn dynamics using limited but representative samples with adaptive strategies. *ISPRS Journal of Photogrammetry and Remote Sensing* 190 (2022), 252–266. https://doi.org/10.1016/j.isprsjprs.2022.06.012
[31] Sitian Xiong, Priscilla Baltezar, Morgan A. Crowley, Michael Cecil, Stefano C. Crema, Eli Baldwin, Jeffrey A. Cardille, and Lyndon Estes. 2022. Probabilistic tracking of annual cropland changes over large, complex agricultural landscapes using Google Earth engine. *Remote Sensing* 14, 19 (2022), 4896.
[32] Jonathan Xu et al. 2024. HarvestNet: A Dataset for Detecting Smallholder Farming Activity Using Harvest Piles and Remote Sensing. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 38. 22438–22446. https://doi.org/10.1609/aaai.v38i20.30251
[33] Leikun Yin et al. 2023. Mapping smallholder cashew plantations to inform sustainable tree crop expansion in Benin. *Remote Sensing of Environment* 295 (2023), 113695. https://doi.org/10.1016/j.rse.2023.113695
[34] Nanshan You et al. 2021. The 10-m crop type maps in Northeast China during 2017–2019. *Scientific data* 8, 1 (2021), 41. https://doi.org/10.1038/s41597-021-00827-9
[35] Liangpei Zhang, Lefei Zhang, and Bo Du. 2016. Deep learning for remote sensing data: A technical tutorial on the state of the art. *IEEE Geoscience and remote sensing magazine* 4, 2 (2016), 22–40. https://doi.org/10.1109/MGRS.2016.2540798
[36] Weichun Zhang, Hongbin Liu, Wei Wu, Linqing Zhan, and Jing Wei. 2020. Mapping rice paddy based on machine learning with Sentinel-2 multi-temporal data: Model comparison and transferability. *Remote Sensing* 12, 10 (2020), 1620. https://doi.org/10.3390/rs12101620
[37] Liheng Zhong, Lina Hu, Le Yu, Peng Gong, and Gregory S Biging. 2016. Automated mapping of soybean and corn using phenology. *ISPRS Journal of Photogrammetry and Remote Sensing* 119 (2016), 151–164. https://doi.org/10.1016/j.isprsjprs.2016.05.014
[38] Liheng Zhong, Le Yu, Xuecao Li, Lina Hu, and Peng Gong. 2016. Rapid corn and soybean mapping in US Corn Belt and neighboring areas. *Scientific reports* 6, 1 (2016), 36240. https://doi.org/10.1038/srep36240