Clark University
Clark Digital Commons

International Development, Community and Environment (IDCE)

Master's Papers

5-2023

Country-Scale Crop Field Mapping for Zambia Using U-Net

Yao-Ting Yao
leannayao@gmail.com

Follow this and additional works at: https://commons.clarku.edu/idce_masters_papers

 Part of the Agricultural Science Commons

Recommended Citation

Yao, Yao-Ting, "Country-Scale Crop Field Mapping for Zambia Using U-Net" (2023). *International Development, Community and Environment (IDCE)*. 261.
https://commons.clarku.edu/idce_masters_papers/261

This Thesis is brought to you for free and open access by the Master's Papers at Clark Digital Commons. It has been accepted for inclusion in International Development, Community and Environment (IDCE) by an authorized administrator of Clark Digital Commons. For more information, please contact larobinson@clarku.edu, cstebbins@clarku.edu.

Clark University
International Development, Community and Environment (IDCE) Master's Papers

5-2023

Country-Scale Crop Field Mapping for Zambia Using U-Net

Yao-Ting Yao
yayao@clarku.edu

Recommended Citation
Yao, Yao-Ting, " Cropland Maps on Planet-Scope Imagery Using Attention U-Net at National Scales in Zambia " (2023). International Development, Community and Environment (IDCE).

Country-Scale Crop Field Mapping for Zambia Using U-Net

Yao-Ting Yao

May 2023

A Master's Paper

Submitted to the faculty of Clark University, Worcester, Massachusetts, in partial fulfillment of the requirements for the degree of Master of Science in the departments of International Development, Community and Environment and the Graduate School of Geography

And accepted on the recommendation of

Lyndon Estes, Chief Instructor

ABSTRACT

Country-Scale Crop Field Mapping for Zambia Using U-Net

Yao-Ting Yao

Cropland mapping is a crucial tool for evaluating food security. Cropland in African countries is expected to expand by at least 140M ha by 2050 to satisfy food demand. Developing accurate, large-area cropland maps of Africa's smallholder agricultural systems is critical. To address this need, we use U-Net, a convolutional neural network, to map cropland for the year 2021 in the Republic of Zambia, a country experiencing rapid agricultural growth. To undertake this work, we first created a labeling platform and workflow protocol on Google Earth Engine (GEE) to collect labels from high spatial-resolution Planet-Scope NICFI Basemaps (4.77m) within grids of 0.005-degree resolution, resulting in labels having dimensions of 200X200 pixels. It provided the growing and off-season as well as false color composite imagery to create labels for training and evaluating the model, collecting a total of 916 over Zambia. To develop the model, we used an existing U-Net model that was trained with 5,377 labels collected across four African countries--Tanzania, Ghana, Nigeria, and the Republic of Congo. The model was trained to recognize three classes (field interior, field boundary, and non-field). This model was then adapted to Zambia by fine-tuning the last 21 layers of the decoder for 300 epochs, using 80% of the 916 collected labels for training and 20% for validation, resulting in an F1 score of 0.64 for the field interior class. The fine-tuned model was used to map predictions for the field interior class for a representative set of 0.05 degree image tiles for the entire country. The results created a 3-meter resolution crop field map Zambia. The accuracy of the map was evaluated with 100 independent reference points, showing an overall, user's, and producer's accuracy of 79, 64, and 91%. The study suggests future opportunities to improve the model by balancing the number of labels among different classes, incorporating slope or DEM data as additional channels to train the model, combining field interior class prediction with field boundary class prediction to reduce false negatives, and applying different CNN architectures in Zambia, such as DenseNet or ResNet.

Keywords: cropland mapping; agriculture; U-Net; Zambia

Lyndon Estes, Ph.D.
Chief Instructor

ACADEMIC HISTORY

Name (in Full): Yao-Ting Yao **Date: May 2023**

Baccalaureate Degree: Bachelor of Science **Date: June 2006**

Source: National Taiwan University

Master Degree: Master of Science **Date: June 2010**

Source: National Taiwan University

Occupation and Academic Connection since date of baccalaureate degree:
Research Assistant, Clark Labs, Clark University.

DEDICATION

I am immensely grateful to my husband, Lucas Lu, whose unwavering support has been instrumental in my journey toward completing my degree and this thesis. Thank you for your understanding and for taking care of Aubrey during the challenging periods of my internship and thesis. Your presence and encouragement have been invaluable.

I would like to extend a special thank you to my academic advisor, Yelena. Her kind words and unwavering support throughout my two-year master's program have been truly inspiring.

ACKNOWLEDGEMENTS

I would like to express my heartfelt gratitude to Professor Lyndon Estes for generously dedicating his time to discuss my ideas, addressing all my questions, and providing invaluable insights. His support and guidance throughout the entire process have been truly inspiring. I am also deeply grateful for his kind and encouraging words, which motivated me to persevere.

I would like to extend a special thanks to Sitian Xiong for his exceptional coding skills and the guidance he provided in coding. His valuable suggestions and tips were instrumental in shaping my thesis. I am truly grateful for his assistance.

Furthermore, I would like to acknowledge the contributions of Sam Khallaghi and the Agricultural Impacts Research Group. Their innovative deep-learning architectures played a pivotal role in the success of this research. Without the groundbreaking work, this project would not have been possible.

TABLE OF CONTENTS

Introduction

Cropland distribution, crop production, and cropland expansion are key factors that affect food security[74], land pressure[75], soil erosion[76], and carbon emissions [77]. Accurate cropland mapping is essential in evaluating food security, as it provides important information, e.g., cropland area, to forecast food production and access potential deficits [78]. To meet the growing demand for food, African countries will need to expand cropland by at least 140 million hectares or increase intensified production in the next decade [1-3, 78]. However, mapping croplands accurately over large areas in Africa is a significant challenge due to limited national survey data and smallholder agricultural systems [5-9, 12, 72-73].

Remote sensing is the most practical method for monitoring agricultural changes, but it also has its challenges, including difficulty in distinguishing smallholder fields from surrounding vegetation [13], limitations of low spatial resolution imagery [13–18], lack of publicly accessible field labels [9, 13, 15-16], computational challenges[13], and frequent cloud cover in tropical and sub-tropical Africa [5]. To address these challenges, we propose the use of Convolutional Neural Networks (CNNs) to learn from high-resolution satellite or aerial imagery and accurately map smallholder-dominated croplands [30, 35-36, 59].

Convolutional Neural Networks (CNNs) have been specifically designed to leverage the spatial information contained in image pixels [59]. In contrast to fully-connected neural networks, e.g. Multilayer Perceptrons (MLP), which treat complex structures as vectors of numbers by flattening the images, disregarding the spatial relationships between pixels, CNNs capture information between neighboring pixels [59]. This ability sets CNNs apart from other pixel-based classifiers, like Random Forest (RF), which overlook the spatial context and consequently produce more 'salt-and-pepper' noise in their results [84]. By employing a convolutional layer, CNNs can effectively extract spatial context, leading to a reduction in the number of parameters and the ability to assign weights to pixels based on their locality [59].

This study focuses on employing a CNN approach to map croplands in the Republic of Zambia for the year 2021. Zambia serves as an ideal study area due to its diverse range of cropland sizes, agricultural production systems, and varying degrees of

cropland expansion [7, 21, 79]. Leveraging high-resolution imagery and utilizing deep learning methods, we aim to accurately map croplands and showcase their effectiveness. The generated cropland map will play a crucial role in estimating the total cropland area in Zambia for 2021. Furthermore, this map serves as a foundational basis for future research endeavors in crop-type mapping. It will provide valuable insights and be incorporated into a vast Africa-wide dataset managed by the Mapping Africa Team.

Background

Study Area

The area of Zambia is 752,618 square kilometers. According to *Climate-Smart Agriculture in Zambia* (2017) [38], land use in Zambia is 5% arable land, 27% permanent meadows, 66% forest area, and 2% other land. The total agriculture area is 23,696,000 ha (32% of the total land area). According to the Crop Forecast Survey by Zambia Statistics Agency, the largest crop planted area in 2017/2018 is utilized for maize (1,392,546 ha, 57.03% of cropped land), followed by groundnuts (284,708ha, 11.66%), soya beans (205,508ha, 8.42%), seed cotton (118,763 ha, 4.86%), and Sunflower (97,851 ha, 4.01%) [10]. The climate of Zambia is sub-tropical. It is characterized by three distinct seasons based on temperature and rainfall: warm rainy season (from November to April) which is the main crop-growth season; cool dry season (from April to August); hot dry season (from August to November) [41].

Agriculture is an important sector in Zambia's economy that creates 22% of GDP and employs 67 % labor force [80]. Zambia's agriculture development program increases agriculture-related services to increase productivity. Zambia's government has continued farming blocks of the land development program to expand agriculture sectors [80]. The policy puts Zambia under rapid crop expansion [5]. There are three types of farmers in Zambia: Small-scale farmers (cultivate less than five ha) are major in Zambia; medium-scale farmers (cultivate between 5 to 20 ha); large-scale commercial farmers (cultivate over 20 ha) use machinery and sell all their production to market [40].

There are three agroecological regions in Zambia (Regions I, II, and III) divided by

rainfall amount, soils, and other climatic characteristics. Region I covers the valley areas in southern, eastern, and western Zambia. The mean annual rainfall in this region is from 600 to 800 mm. The growing season is relatively shorter than other regions (80-120 days). Soils constrain crop production. Small-scale farmers are predominant in this area. Region II is range from most of Central, Southern, Eastern, and Lusaka provinces. This region has fertile soils and most of the country's commercial farms. The mean annual rainfall is 800-1000 mm. The growing season is 100-140 days. Region III is in the Northern Luapula Copper belt, Northwestern provinces, and some parts of the Central province. This region receives the highest mean annual rainfall (over 1000 mm), and the growing season is 120-150 days. Small-scale farming predominates in this region. Farmers use very low input and shifting cultivation techniques. [40]. Frequent cloud cover occurs during the growing season [5, 7, 21], especially in Region III.

Research Review

A large number of studies focused on mapping croplands in Africa have been undertaken, including several in Zambia. Sweeney et al. (2015) used spectral signatures to develop land cover classification in Zambia's Southern Province [7]. The workflow combined statistical clustering, supervised classification, proportional sampling, and targeted error detection with a probabilistic reclassification technique. They applied a clustering algorithm (ISODATA) in Erdas IMAGINE with Landsat 5 Thematic Mapper (TM) data for the pre-season of 2008 and the harvest season of 2009 to create a training dataset. All images are grouped into ten clusters and randomly sampled at 75 points. A total sample of training locations is 750 points. 498 were labeled with suitable land cover classes: (1) forest; (2) cropland; (3) savanna; (4) settlement; and (5) water. Spectral signatures were extracted and produced mean subgroup signatures. 477 validation random sampling of the 25-category thematic map to generate validation points. Classification error was evaluated, and reclassification was iterative. The author assessed classification error and use logistic regression to estimate the probability. The overall map accuracy is 88.18%. Omission error within the cropland class is 12.11% and commission error is 9.76% [7].

Xiong, S., et al. (2022) used Bayesian Updating of Land Cover Unsupervised (BULC-

U) algorithm to detect the locations and dates of cropland expansion [5]. They applied this approach for a time series of unsupervised classifications developed from Landsat 5, 7, 8, Sentinel-1, and ALOS PALSAR within 1476 tiles and mapped annual cropland change from 2000 to 2015 in Zambia. The result shows active cropland expansion between 2000 to 2015 in Zambia, especially in the Southern, Central, and Eastern provinces.

Potapov, P. et al. (2022) created global cropland map at 4-year intervals (2000–2003, 2004–2007, 2008–2011, 2012–2015 and 2016–2019) using 30 m spatial resolution Landsat satellite [72]. The Global cropland mapping included three stages. Each stage used bagged decision tree that used 924 tiles with class presence and absence as the dependent variables, and a set of multitemporal metrics converted from Landsat imagery as independent variables. The results created a cropland probability layer and then used a threshold of 0.5 to obtain a cropland map.

Lobell, D.B. and Azzari, G. (2017) use Random Forest classifiers to do land use classification at the national level of Zambia [6]. They used the Landsat archive within Google Earth Engine to download Landsat 7 TM, Landsat 7 ETM+, and Landsat 8 OLI sensors intersecting the boundaries of Zambia between 2012 and 2015. The authors trained a Random Forest classifier on two types of Landsat composites. (1) A season-based composites: four seasonal composites included all reflective bands, NDVI (Normalized Difference Vegetation Index), and GCVI (Green Chlorophyll Vegetation Index). (2) A metrics composite: 0.1, 0.25, 0.5, 0.75, and 0.90 quantile composite for all the reflective bands, including NDVI and GCVI, totaling 40 metrics.

They extracted per-band pixel values of both composites from randomly selected training points and used the resulting 30% of points data to train Random Forest classifiers and 70% to test on GEE. They classified Six Land use classes: 1) rainfed crops (RFC), 2) irrigated crops (IRC), 3) open-canopy natural vegetation (OCN), 4) close-canopy natural vegetation (CCN), 5) swamp natural vegetation (SWN), and 6) urban and suburban areas (URB). In the result, 89% accuracies were achieved for the season- and metric-based approaches for individual classes, with 93% and 94% accuracy for distinguishing cropland from non-cropland [6].

Related to cropland mapping, there is an additional body of work that focuses on

methods for delineating individual crop fields. Lesiv, M. et al. (2019) pointed out two approaches for field boundary mapping [42]. Remote sensing is a commonly used method for field boundary mapping, but it has yet to be fully implemented at a global scale. Additionally, census data is often used to estimate the average size of fields at a subnational level, but the spatial resolution of this method is typically quite coarse. The authors allocated the sample units within four existing cropland maps: a cropland layer derived from Globe land 30 at a 30 m resolution [43]; a cropland layer derived from the ESA CCI LC map at a 300 m resolution for 2015 [44]; the unified cropland layer at a 250 m resolution [45]; the IIASA-IFPRI hybrid cropland layer at a 1 km resolution [46]. They used the nearest-neighbor method to create a global field size map and evaluate the accuracy by compared with the control sample. When the fields identified by the experts matched a pixel value on the field size map, this classification was true. The results show that smallholder farms occupy up to 40% of agricultural areas globally [42].

Wang et al (2022) employed U-Net model in Landsat composite imagery over the US Midwest using two types of weak supervision labels (pixel labels and image labels) to create cropland segmentation [81]. They deployed the U-Net result to segmentation and applied it to other pixel-level machine learning algorithms, e.g., logistic regression, support vector machine, and Random Forest. This approach demonstrated U-Nets trained on weak labels with as few as 100 labels and the result as superior classification would allow pixel base labels to be obtained from image labels. The result also showed greater spatial coherence in their predictions.

Waldner and Diakogiannis (2020) extracted field boundaries from Sentinel-2 data using a convolutional neural network [82]. They used ResU-Net, a convolutional neural network with a fully connected U-Net backbone to extract field boundaries.

Estes, L., et al. (2022*)* addressed two problems of field boundary mapping [13]: 1) the spatial and temporal mismatch between satellite sensors and smallholder fields, and 2) lack of high-quality labels for training and assessing the classifiers. The authors create two Planet-Scope cloud-free image composites of the growing season and dry season of a year in Ghana. The seasonal contrast between the two composites helped improve classification accuracy. They also created a platform to assess and minimize label error and used it to train a Random Forest classifier to identify the most effective

training sample through prediction uncertainty. They used cropland probabilities to create crop field boundaries. Cropland probability and field boundary maps had accuracies of 88 and 86.7%; user's accuracies are 61.2 and 78.9%, and producer's accuracies are 67.3 and 58.2%.

Estes, L. et al. (2022) [47] used an updated method based on U-Net (a convolutional neural network [48]) to create an improved set of field boundary maps for Ghana for the year 2018. They used a version of the model trained on 2018 Planet-Scope imagery with 4593 labels with collected across Ghana, Tanzania, and the Republic of Congo. They trained the model with three classes (field interior, field boundary, and non-field). They froze the weights on the first 58 layers of the model and updated parameters on the remaining layers over 15 epochs. The accuracy of this initial model was 86.1% for the field interior class. The resulting refined model was 83.1% for the field interior class, with true and false positive rates of 61.9 and 11%.

CNN

Convolutional Neural Networks (CNNs) are a specialized form of neural networks that have been designed for image processing and computer vision applications [66]. They comprise multiple elements, encompassing convolutional layers, activation function layers (e.g., ReLU), batch normalization layers (e.g., BN), pooling layers, and a final classifier (e.g., sigmoid or softmax).

In the sliding window approach of CNNs, the convolutional layer selects small windows, known as the convolution kernel, of a given size and weight [59]. This allows the CNN to capture local information and preserve the spatial structure of the pixels by computing corresponding hidden representations within a small neighborhood [59]. As a result, a feature map is generated. However, reducing the number of parameters through the convolutional layer can potentially limit the model's expressive power. To overcome this limitation, padding the image with zeros around its boundary provides sufficient space for shifting the kernel. Furthermore, increasing the number of channels can enhance the complexity of the model [59].

The activation function layer determines whether a neuron should be activated based on a weighted sum, which includes the addition of a bias value [59]. The rectified

linear unit (ReLU) function is commonly used as the activation function in CNNs [67]. ReLU retains only positive elements and sets the activations of negative values to 0 [59].

Batch normalization is applied after the convolutional layer and before the activation function. It normalizes the inputs by subtracting the mean and dividing by the standard deviation of the current minibatch. This process also incorporates a scale coefficient and an offset to recover any lost degrees of freedom [59].Pooling layers mitigate the sensitivity of convolutional layers to location and downsampled spatial representations [59]. They calculate either the maximum or average value of elements in a pooling window. For example, the U-Net architecture uses Max-pooling, which computes the maximum value in the window from upper-left to bottom-right at each location [59].

U-Net

U-Net is one of the CNNs architectures for image segmentation. It comprises an encoder (U-shape construction path) and a decoder (expansion path) structure, with the spatial scale subsequently reduced after consecutive pooling operations and increased in a contracting path [49]. This study utilized the U-Net architecture from DeepLearner, a deep learning package developed and maintained by Mapping Africa Team since 2021. The package provides Pytorch implementation for multiple CNN architectures including U-Net.

In this research, the U-Net architecture contains six encoder and five decoder layers in its contraction path (Figure 1). To form the encoder, the VGG-16 [92] architecture was adapted as the backbone. The VGG16 architecture consists of thirteen convolutions with kernels with padding of 1, five max-pooling layers with a stride of 2, and three dense layers. VGG-16 consists of thirteen convolutional layers with 1-pixel padding, five max-pooling layers with a stride of 2, and three dense layers. The first two VGG blocks contain two convolutions and a max-pooling layer, while the third, fourth, and fifth blocks consist of three convolutions and a max-pooling layer.

For this research, VGG blocks are adapted to the convolution blocks. Each encoder step includes two 3 × 3 convolutional layers followed by batch normalization and a

ReLU activation function. The spatial dimensions of the features are reduced using a 2 × 2 max-pooling layer with a stride of two. To incorporate both low-level and high-level information, skip connections are employed. These connections concatenate information derived from the convolutional layers, enabling the network to leverage both types of information and produce more accurate segmentation results [69]. This approach is particularly beneficial when training data is limited, as it helps produce accurate prediction maps [50]. If skip connections are used, the high-level semantic information from the encoder is concatenated with the low-level information from the decoder.

During the decoding process, transposed convolution layers are utilized to up-sample the image. These layers increase the spatial dimensions of intermediate feature maps, which were previously reduced by the convolutional layer [59]. Each element in the input tensor is multiplied by the kernel and summed over to produce the output [59]. Strides are specified for intermediate results, allowing the output to increase both the height and width of intermediate tensors [59]. In this research, a transposed convolution with a 2 × 2 kernel and a stride of two is used for each decoder layer. Finally, a Softmax classifier is employed to generate the three-class result.

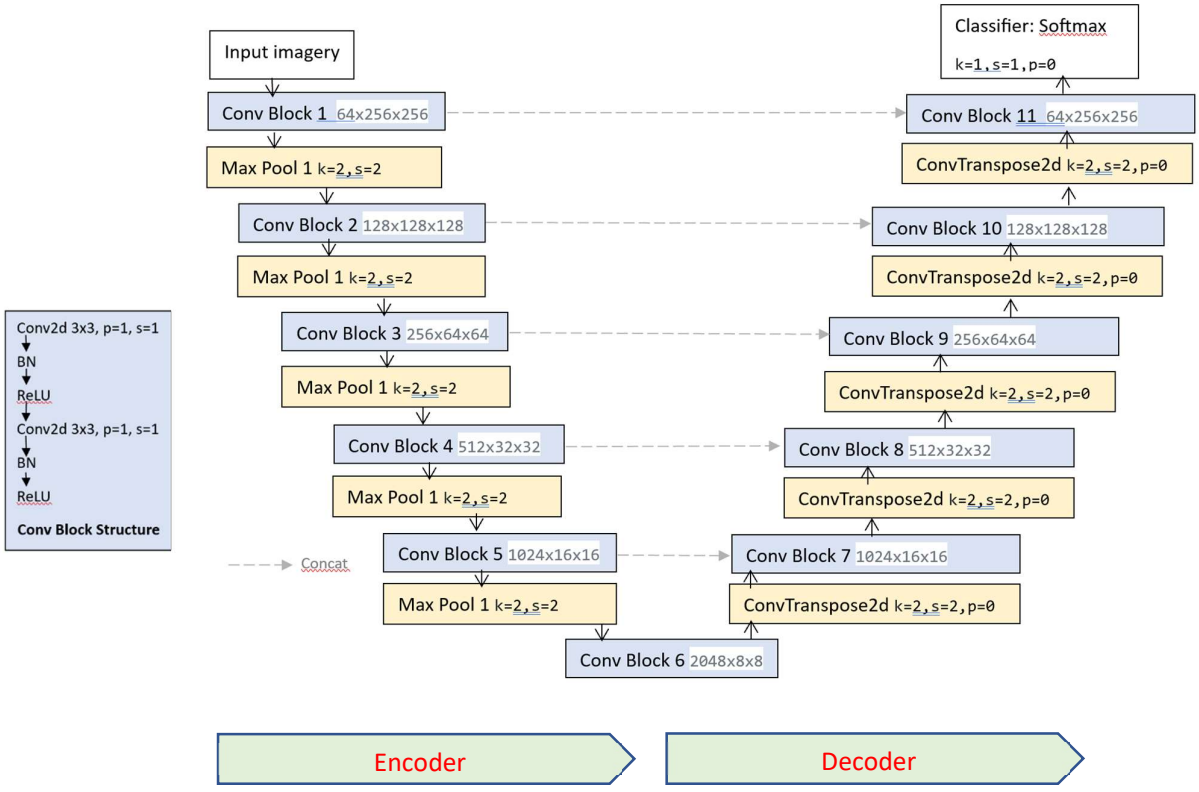


Figure 1 U-Net architecture in this research

Research Objective

The primary objective of this research is to evaluate the performance of the U-Net deep learning model in predicting cropland types across a diverse range of agricultural landscapes, including small-scale swidden croplands and large commercial farmland. The aim is to compare the predictive capabilities of U-Net with those of RF, another commonly used machine learning classifier. The research also considers the availability of limited training labels, reflecting real-world constraints. This constraint poses a challenge in terms of data production, particularly in deploying U-Net effectively on a cloud computing platform.

By addressing these research questions, the study aims to develop a prototype cropland map specifically for Zambia. This map will serve as a valuable tool for assessing and monitoring food security in the region. Furthermore, this map will also be incorporated into a vast Africa-wide dataset managed by the Mapping Africa Team, which covers various regions across the African continent.

Data and Methodology

Data

1. Planet-Scope NICFI Basemaps

This research made use of high spatial resolution (4.77m) Planet-Scope mosaic basemaps created by Norway's International Climate & Forests Initiative (NICFI). These basemaps, designed specifically for monitoring tropical forests, encompass Tropical Africa, Tropical Asia, and Tropical Americas, covering the region between 30N and 30S [88]. The basemaps consist of both biannual data (from December 2015 to August 2020) and monthly collections (from September 2020 onwards) [88]. They are derived from daily Planet-Scope images, which capture the clearest scenes available on a daily basis, and are subsequently merged into a monthly mosaic [88]. The basemaps are provided in a 4-band multispectral format, including Blue, Green, Red, and NIR bands.

The research calculated the mean of monthly basemaps for July, August, and September to create a full image catalog for the entire country that represented the dry

season for Zambia. Imagery from the dry season is least contaminated by cloud, and field boundaries typically are still clearly visible at this time of year.

2. Global Cropland Map

The Global Land Analysis and Discovery (GLAD) laboratory at the University of Maryland's Department of Geographical Sciences developed a comprehensive global cropland map [72]. This dataset offers a consistent representation of cropland extent over time, with a spatial resolution of 30 meters. The mapping process relied on the use of Landsat satellite data from 2000 to 2019, which was consistently processed and archived. Specifically, this research utilized the Global_cropland_NE_2019 dataset, which includes crop and non-crop classes. The training and validation sample points were extracted from this cropland map. These sample points were then utilized in the RF model, and the results were compared with predictions generated by the U-Net model.

Methods

This study adapted the methods proposed by Estes et al. (2021) [13] and Estes et al. (2022) [19] to develop an approach for generating cropland maps. Figure 2 illustrates the conceptual diagram of the research methodology.

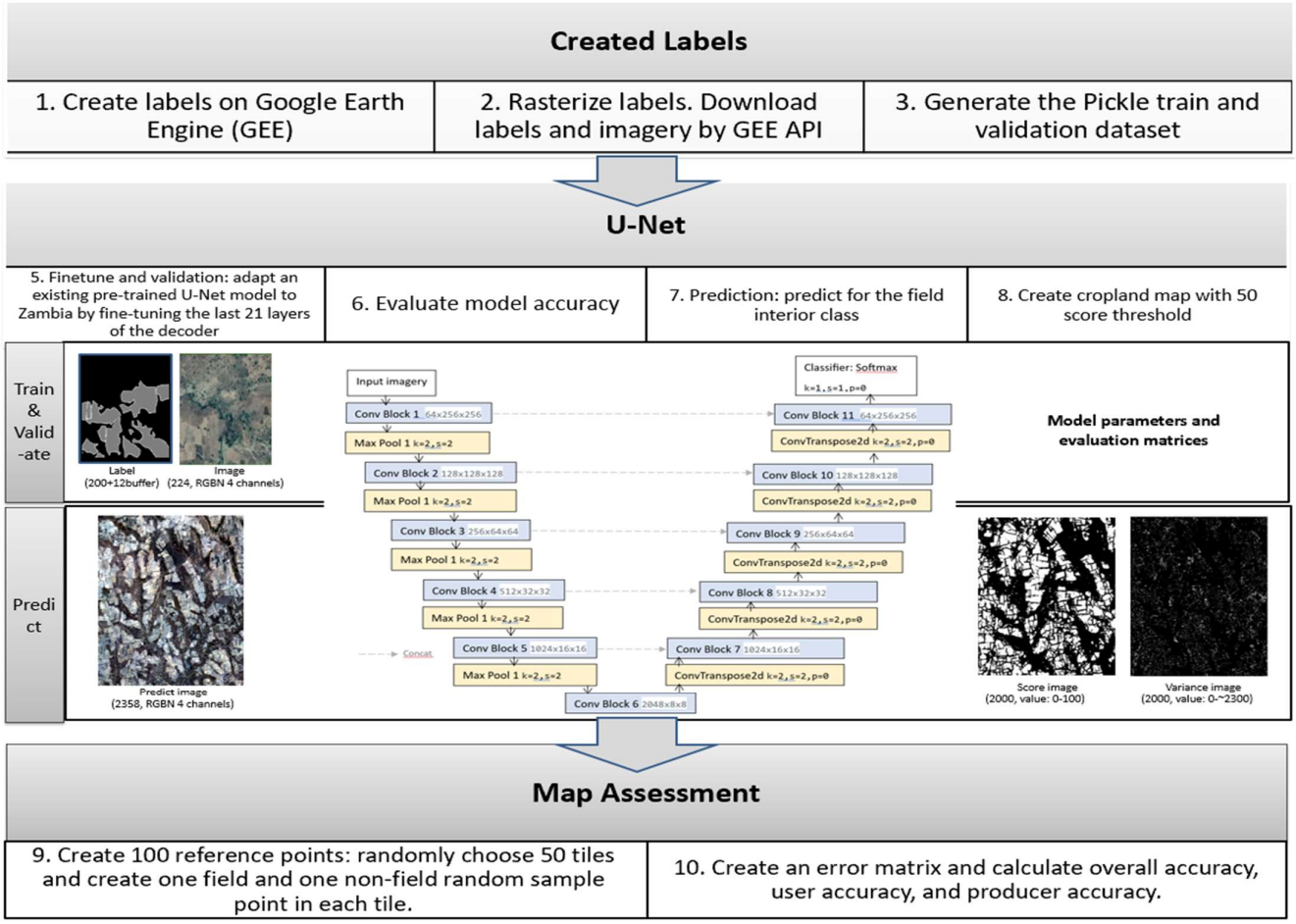


Figure 2: Conceptual diagram

Create a labeling platform and workflow protocol on Google Earth Engine (GEE)

To cover the entire region of Zambia, a grid of 0.005 degrees was generated, and a catalog was created using R script to randomly select 1800 grid cells across three agriculture zones based on Global Food Security Support Analysis Data at 30 m (GFSAD) to ensure labels with fields. To create labels for training and evaluating the model, a labeling platform and workflow protocol were established on Google Earth Engine (GEE), providing 0.005 degrees high spatial resolution Planet-Scope NICFI Basemaps grids of the growing and off-season, as well as false-color composite imagery (Figure 3). A total of 916 labels were collected over Zambia (Figure 4).

Figure 3 Labeling interface on GEE

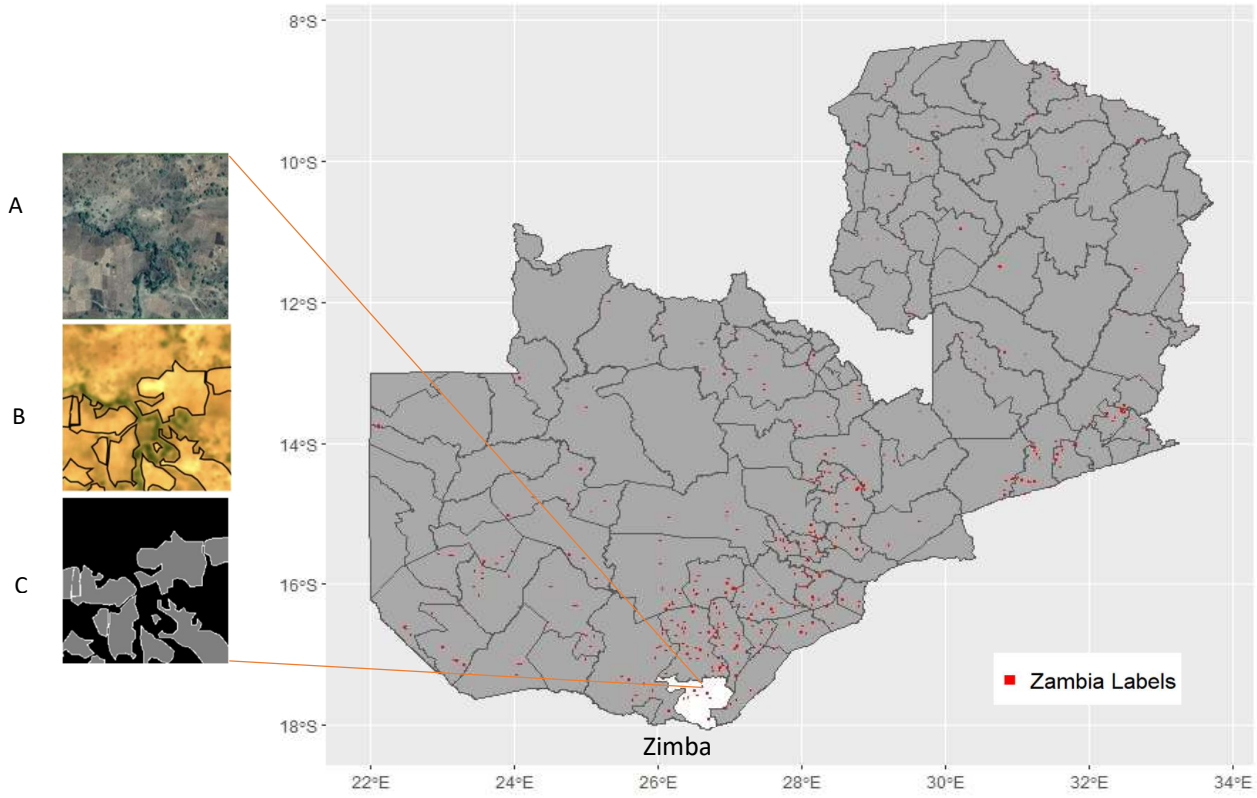


Figure 4 Zambia Labels Created by this research. A is a Planet-Scope NICFI Basemaps. B is a label feature class created on GEE. C is a raster label.

The Google Earth Engine (GEE) API was used to rasterize labels into three distinct classes: non-field (class 0), field interior (class 1), and field boundary (class 2). The cubic resample method was then used to resample each label to the output resolution. The label size dimensions were set to 200 by 200, or 0.000025° (<3 m), with a 12-pixel buffer padding, while the Planet-Scope NICFI Basemaps was set to 224 by 224 divisible by 32 due to five layers of down-sampling and up-sampling to process results in the same resolution for input and output imagery [48]. To avoid losing pixels on the perimeter of the image, especially in the corners of the pixel when applying convolutional layers in any CNN model [59], adding 12 extra pixels with zero value of padding around the boundary were added to each labeled image. Additionally, an imagery downloader was also created using the same GEE API to download Planet-Scope NICFI Basemaps chips for training. The prediction imagery

was 0.05-degree image tiles. 179 pixels buffer added in each tile using R script to generate the tile grid shapefile covering the whole of Zambia and upload it to GEE. The GEE API was then used to download all the 24,952 tiles and resample them into 2358 by 2358 dimensions.

Two labelers, Ph.D. student Xiong, S. and I, created the labels in this research. The labeling workflow consisted of two steps, label score assessment, and recreating labels, to ensure label quality. In the first round of label creation, each labeler accesses each label on their own and gave a score between 1 to 10 after labeling, with 10 being the most accurate. If the label score was lower than 5, I reviewed and recreated the labels in the second round to improve the label's accuracy. Non-field labels were selected in the area with high false positives. The total number of non-field labels needed to be between 10% to 15% of total labels to improve overall model accuracy based on empirical experience from the Agricultural Impacts Research Group. After each round of label creation, the model was trained, and several tiles were predicted to see if there was a need to create new labels.

Finetuning U-Net for Zambia

To develop the model, we used an existing U-Net model that was trained with 5,377 labels collected across four African countries--Tanzania, Ghana, Nigeria, and the Republic of Congo. The model was trained to recognize three classes (field interior, field boundary, and non-field). This model was then adapted to Zambia by freezing 58 layers in which the layer's weights of the pre-trained model remain frozen and fine-tuning the last 21 layers of the decoder for 300 epochs.

Through set random seeds for samples of train and validation when making a catalog in R script, this research used 720 collected labels (80% of total labels) for training and 196 collected labels (20% of total labels) for validation. Prior to training, it was necessary to generate training and validation datasets through the data loader, which applied the same series of image augmentations that were used in the previous model, which included vertical flip, horizontal flip, rotation with a range from -90 to 90 degree, resize and shift brightness. Image data augmentations help to increase training dataset size and reduce model overfitting [89-91]. To facilitate dataset

loading and reduce the processing time, the processed samples were converted to pickle files [70]. The Nesterov accelerated gradient optimizer was chosen to speed up the weights and biases adjustment process [62]. The following hyper-parameters were used during training, validation, evaluation, and prediction:

Train batch size	32
Validation batch size	2
Epochs	300
Dropout rate	0.15 (for model fitting); 0.1 (for prediction)
Learning rate	0.01 (using polynomial learning rate decay as learning rate policy)
Prediction patch size	250
Prediction buffer	179
Composite buffer	179
Prediction batch	2
Shrink pixels	54
Number of MC trials	10

Table 1 Hyper-parameters

After model fitting, evaluating the model helps us generate a series of model performance scores. F1 score is the combination of precision and recall using the harmonic mean [83]. The larger value of F1 score implies larger value of precision and recall. The formula of precision, recall and F1 are as follows. TP is true positive; FP is false positive; FN is false negative.

$$precision = TP/(TP+FP)$$

$$recall = TP/(TP+FN)$$

$$F1=2* precision*recall/(precision+recall)$$

During prediction, the input image of 2358 by 2358 pixels was divided into smaller patches of 250 pixels. A prediction buffer of 179 pixels was used to ensure that the output image had smooth transitions between neighboring patches. The composite buffer of 179 pixels was used to ensure that the final output image had smooth transitions between all the patches. According to Agricultural Impacts Research Group, setting up a buffer helped us to minimize edge effects. When a patch was processed, the model predicted an additional buffer of 179 pixels around the edges of the patch. This buffer ensured that the model captured the context around the edges of the patch and avoided any sharp edges in the output image where the patches were stitched together. After removing this buffer, the final prediction dimension was 2000 by 2000 pixels.

To enhance the robustness of the model's predictions, this research incorporated the use of Monte Carlo dropout (MC) during the prediction phase [61]. Dropout is a technique employed during training, where certain neurons are randomly deactivated based on a predefined dropout rate, controlling the probability of each neuron being turned off [85, 86]. By utilizing dropout, the model can mitigate overfitting by disrupting the dependence of each layer on specific activation patterns from the previous layer [59]. However, during prediction, dropout is turned off to leverage the entirety of the trained neurons and connections, which can introduce uncertainty in the results [86].

To address this uncertainty and obtain a measure of model uncertainty, MC dropout is employed during prediction [87]. This approach involves running the model multiple times with dropout enabled and subsequently averaging the outcomes to obtain more reliable and robust predictions [86]. The number of MC trials determines how many times the model is run with dropout enabled and the results are averaged [61]. After testing various MC dropout rates, such as 0, 0.05, 0.1, 0.2, and 0.3, this research established a dropout rate of 0.1 and performed 10 MC trials to achieve better prediction result. Following the prediction phase, a cropland probability map was generated. A threshold probability of 0.5 was applied to produce the final cropland map.

Creating Random Forest Cropland Map

In order to compare the performance of the U-Net model with the RF classifier, this research employed the GEE smileRandomForest module to create an additional cropland map specifically for Zambia. The training and validation of this classifier utilized six variables, namely R, G, B, NIR, NDVI, and GCVI. Similar to the approach used for the U-Net model, the Planet-Scope NICFI Basemaps was extracted using the R, B, G, and NIR bands from the mean values of off-season imagery collected between July 1, 2021, and September 30, 2021. Additionally, the Normalized Difference Vegetation Index (NDVI), calculated as (NIR - Red) / (NIR + Red), and the green chlorophyll vegetation index (GCVI), calculated as (NIR/Green) - 1, were included as variables.

The point labels used in this study were generated from the Global Cropland Map of 2019, where a value of 1 indicates cropland and 0 represents non-crop areas. In comparison to the point labels generated from 916 labels created specifically for this research, the point labels derived from the Global Cropland Map across the entire country exhibited superior prediction results (Figure 5).

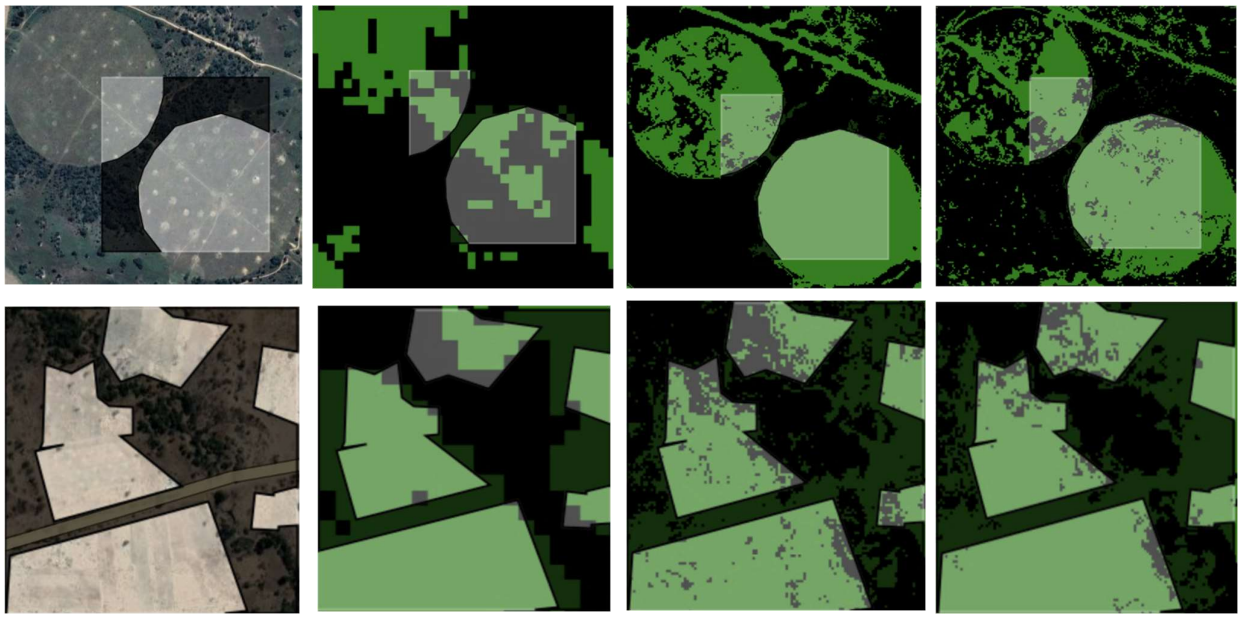


Figure 5 RF Prediction with different types of point labels. A are true-color composites of Planet-Scope NICFI Basemaps with labels created from this research (White is field; black is non-field), B are 2019 Global Cropland Map, C are RF prediction with point labels generated from 916 labels, and D are RF prediction with point labels generated from 2019 Global Cropland Map.

To obtain representative sampling points, a total of 1000 points were collected throughout Zambia, based on the 2019 Global Cropland Map. These points were stratified using the GEE stratifiedSample module, ensuring that 500 points corresponded to crop areas while the remaining 500 points represented non-crop areas. The stratifiedSample module randomly assigned column values to split the dataset, with 80% of the points allocated to the training dataset (column value <= 0.8), and the remaining 20% used for the validation dataset (column value > 0.8). The validation dataset played a crucial role in assessing the performance of the RF classifier, allowing for the calculation of accuracy metrics.

Map assessment

Stehman and Foody (2019) provide guidelines for conducting an independent assessment of the categorical accuracy of final maps [55]. Estes et al. (2022) used an

independently collected point sample to calculate overall accuracy for each map, class-wise User's and Producer's accuracy, and the 95% confidence intervals for each accuracy measure across different zones and the entire country [13, 56-58].

To perform the accuracy assessment, this research employed a random sampling strategy as follows:

1. Randomly select 50 tiles and ensure that each tile has both field and non-field classes.
2. Randomly select one field and one non-field reference point in each tile.
3. Review Planet-Scope NICFI Basemaps under each point to determine whether each point represents a field or not.
4. Compare the reference points to the prediction maps and create a confusion matrix by counting and calculating overall accuracy, user's accuracy, and producer's accuracy.

Accuracy rates range from 0 to 1, where 1 represents 100% agreement between the sampled value from map and the label. The user's accuracy indicates false positives error (same as specificity) or errors of commission, where pixels are wrongly classified as a known class when they should have been classified as another class. The producer's accuracy, on the other hand, indicates false negatives (same as sensitivity or recall) or errors of omission [60]. The producer's accuracy reflects how well the classification results meet the prediction expectations [60].

This implementation was written in Python using the *os* library to create the list and the *random* package to randomly select 50 prediction tiles from the list. To create the random reference points, the GEE ee.FeatureCollection.randomPoints module was used, with seed parameters set. The numbers of counts in the confusion matrix were calculated, as well as overall accuracy, user's accuracy, and producer's accuracy.

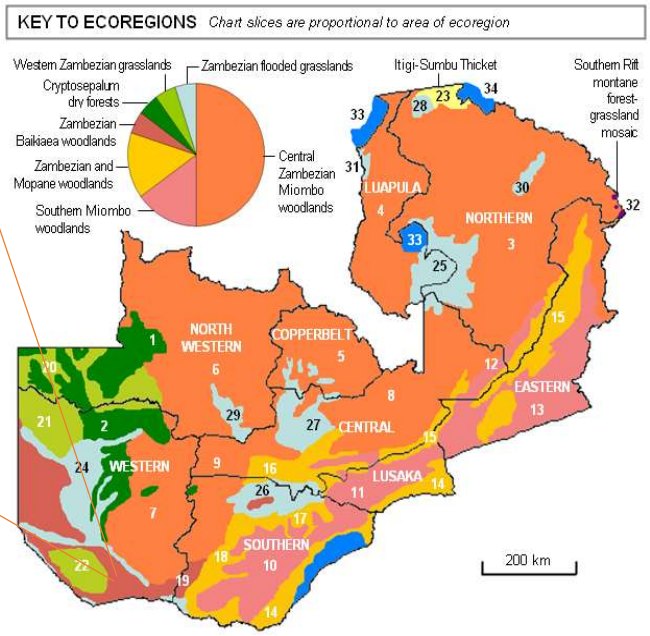
Results

The study generated a partial map of Zambia's annual croplands regarding the research time limitation, comprising 7,008 tiles covering an area of 195,001km^2, mostly in northern, southern and central Zambia, which represents 28% of the entire nation's land area .

Label creation and quality assessment

Two labelers created the labels in this research, with Ph.D. student Xiong, S. creating 301 labels and me creating 615 labels. The initial model results showed high false positives in the wetland area, and large roads were classified as fields. A second round of labeling was therefore undertaken, during which I created labels on the wetland and riverbank with fields and non-fields, as well as small farmland with a larger road system area. Additionally, non-field labels were selected in areas with high false positives, especially in the western central Zambezian Miombo woodlands (figure 6).

A



B

Figure 6 illustrates the higher false positives that occurred in the Miombo woodland area. A shows the ecoregions in Zambia, as depicted by Malambo, F. M., and Syampungani, S. (2008) [62]. B is a Planet-Scope Basemap of the flat sandy Zambezian Miombo woodland where we created non-field labels to improve the accuracy of the model.

Figure 6 depicts the Zambezian Miombo woodlands around the Zambezi Floodplain, one of Africa's great wetlands, on the Zambezi River in the Western Province of Zambia [64]. The shape of the floodplain area in the flat sandy Zambezian Miombo woodlands confused the model in the first round of prediction.

Model Evaluation

Following the fitting of the model, validation samples were employed to assess its performance. The initial global model achieved an F1 score of 0.66 for the field interior class. Subsequently, the U-Net model was fine-tuned using the labels specific to Zambia for 300 epochs, and the evaluation scores were computed and presented in Table 2.

	accuracy	precision	recall	F1-score	IoU	AUC
class_1: field interior	0.846049	0.706066	0.591981	0.64401	0.474937	0.84332
class_2: field boundary	0.924468	0.133479	0.419668	0.202538	0.11268	0.772429

Table 2: Evaluation metrics for U-Net model performance

In the case of the field interior class (Table 2), the model has relatively higher overall accuracy (0.84), AUC (0.84), and precision (0.70) metrics. However, the lower recall (0.59) and higher precision (0.70) values indicate more false negatives than false positive errors. The F1 score (0.64) is lower than the global model caused by the lower recall scores. Additionally, the Intersection over Union (IoU) metric showed

that the predicted and target chips only had a 47% overlap, indicating room for improvement in the model's ability to accurately detect in the field interior class.

In RF model evaluation, validation overall accuracy is 0.837837. The variable importance chart indicated Planet-Scope NICFI Basemaps R band is the most important variable for cropland mapping (Figure 7).

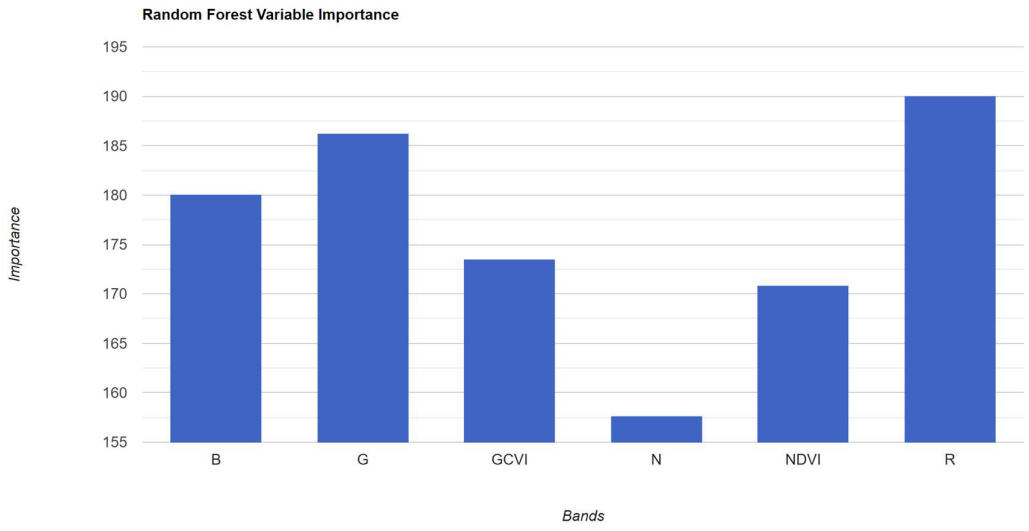


Figure 7 Variable Importance Chart

Zambia's Croplands

The cropland map generated by the U-Net and RF models provides valuable information regarding the distribution of fields in Zambia. It reveals that small farmlands are clustered in the southern region, while large commercial farmlands are predominantly located in central Zambia. Additionally, slash-and-burn agriculture is prevalent in the northern and western parts of the country (Figure 9). This mapping technique also enables the estimation of the total cropland area in Zambia for the year 2021.

We used Zimba District as an example, which is located in southern Zambia and encompasses a total area of 5,234 km2. This district is recognized as the primary agricultural hub in the region. By employing the per-pixel mapping approach, we estimated that the extent of cropland in Zimba District was approximately 1,115 km2, accounting for 21.3% of the district's total area.

To establish a comparative analysis, we referred to the 2019 Global Cropland Map, which previously estimated the cropland extent in Zimba District to be 1,542 km2, representing 29.4% of the district's total area. Conversely, when employing the RF model, our cropland map estimated the cropland extent in Zimba District to be 3,799 km2, accounting for the same 72.5.% of the district's total area (Figure 8).

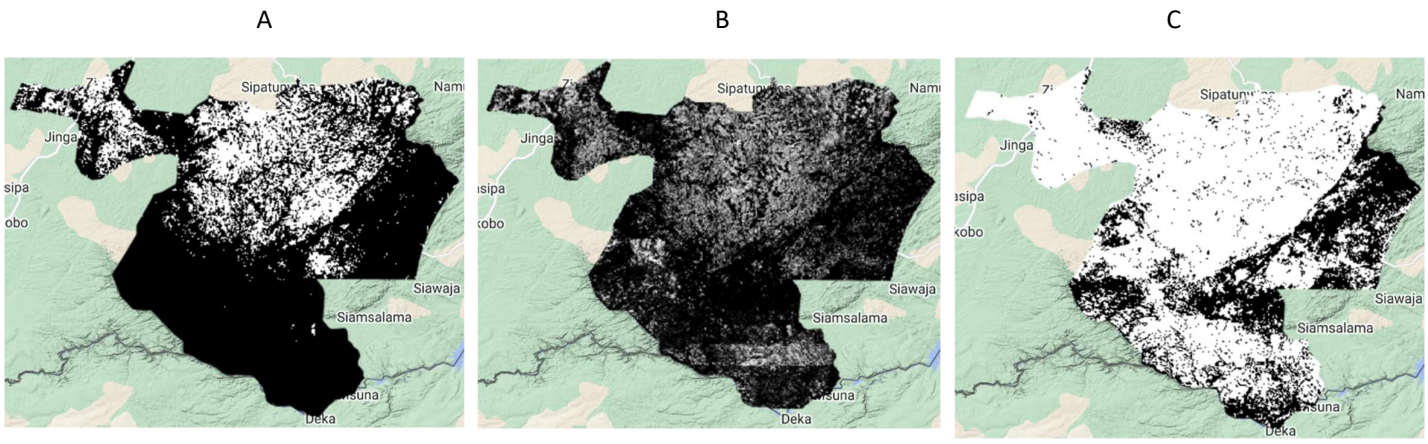
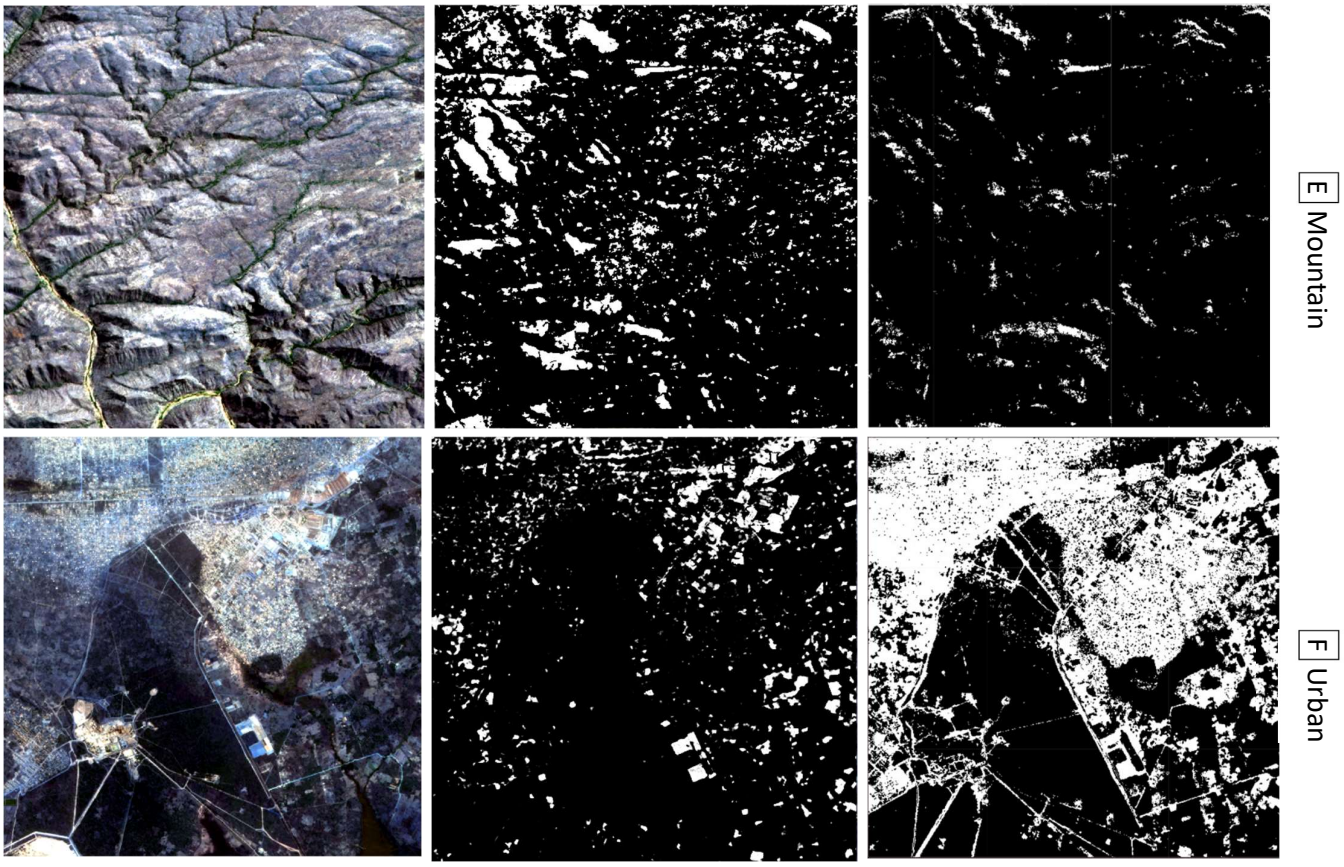


Figure 8 Cropland maps in Zimba District. A is the 2019 Global Cropland Map, B is the U-Net prediction map, and C is the RF prediction map.

Planet-Scope Nicfi Basemaps
U-Net
RF
A Small holder fields
B Commercial farmland
C Slash-and-burn
D Wetland



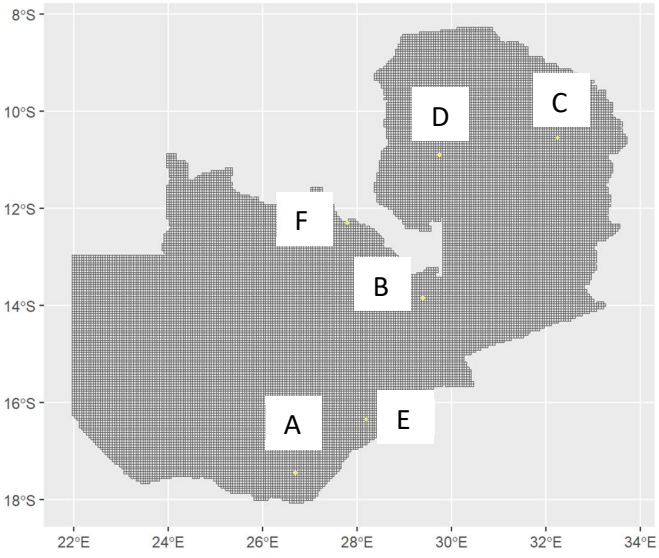


Figure 9 Cropland map examples from U-Net and RF. Each row represents an example tile. The left column is the True Color Image from Planet-Scope NICFI Basemaps, the middle is the Predicted Cropland Mask from U-Net, and the right column is RF prediction. White in the masks indicates crop, and black indicates non-crop.

Map assessment

The map assessment of U-Net and RF cropland maps was conducted based on the reference points depicted in Figure 10. U-Net prediction achieved an overall accuracy of 79% (Table 3A). Specifically, for the field class, the accuracies ranged between 64% (producer's) and 91% (user's), indicating significant omission errors in the model's classification results for this class. The lower producer's value suggests a higher incidence of false negatives within the cropland classified results. Conversely, for the non-field class, the accuracies ranged between 72% (user's) and 94% (producer's), suggesting higher commission errors in the model's classification for this class. The lower user's accuracy value implies a high omission error in the cropland class.

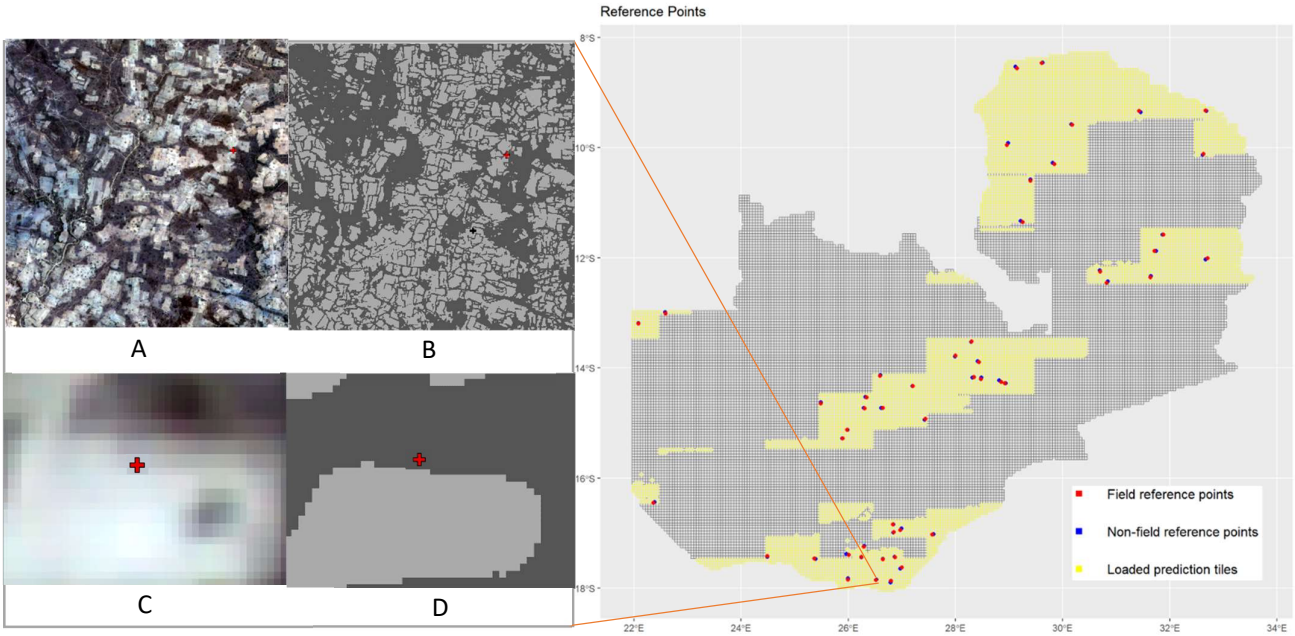


Figure 10 Map assessment points using tile ZM938396 as an example. The map shows all loaded prediction tiles, 50 field reference points, and 50 non-field reference points. A is true-color composite of Planet-Scope NICFI Basemaps with reference points. B displays the U-Net prediction result. C is zoomed-in view enables a closer examination of the actual field reference point on Planet-Scope NICFI Basemaps. In D section, the prediction result indicates a non-field classification, which represents a false negative occurrence commonly observed around field boundaries.

In the case of RF prediction, the overall accuracy achieved was 71% (Table 3B). For the field class, the accuracies ranged between 68% (producer's) and 72% (user's), indicating significant omission and commission errors in the model's classification for this class. The lower producer's and user's values suggest a higher incidence of false negatives and false positives within the cropland classified results. Similarly, for the non-field class, the accuracies ranged between 72% (user's) and 74% (producer's), indicating higher omission and commission errors and implying higher false-negative and false-positive values within the non-cropland classification.

A. U-Net Confusion Matrix

Reference Data					User Accuracy
		field	non-field	total	
Prediction Data	field	32	3	35	91%
	non-field	18	47	65	72%
	total	50	50	100	
Producer Accuracy		64%	94%		**Overall Accuracy:** 79%

B. RF Confusion Matrix

Reference Data					User Accuracy
		field	non-field	total	
Prediction Data	field	34	13	47	72%
	non-field	16	37	53	72%
	total	50	50	100	
Producer Accuracy		68%	74%		**Overall Accuracy:** 71%

Table 3 Confusion matrix

Higher false negatives of the cropland class be observed at the edge of the field, indicating that the field edge had been classified as a non-field class in prediction.

Discussion

Limitation of label creation

The labeling task was one of the major limiting factors, as determining croplands from non-croplands can be difficult within 5 m resolution imagery, especially in smallholder-dominated agriculture systems (Estes et al, 2022). The distinction between active and inactive fields was largely dependent on the discretion of individuals. The interpretation was particularly challenging where croplands and abandoned croplands had only slightly different reflectance. However, the score was created by individual labelers, and the label assessment process should establish criteria to ensure that the score's standards are consistent between different labelers.

Improved Detection of Individual Fields Using U-Net Compared to Random Forest

Our study findings showcase the efficacy of employing U-Net in conjunction with high-resolution imagery for the purpose of mapping croplands. The overall accuracy of U-Net (79%) is higher than RF's overall accuracy (71%). In the field interior class, the user accuracy of U-Net (91%) is much higher than RF (72%) which indicates that more false positives happen in the cropland map of RF than in U-Net, especially in urban areas and road facilities (Figure8). The producer accuracy of U-Net (64%) is a little lower than RF (68%) which indicates that more false negatives happen in both the cropland map of RF and of U-Net, especially around field boundaries (Figure9).

Our approach outperformed other methods such as RF, as shown in Figure 8. Specifically, our results indicate that our model effectively delineates field boundaries in single and unmixing agriculture types region, including smallholder fields, large commercial farmland, and slash-and-burn agriculture. These findings highlight the potential of U-net for advancing the accuracy and efficiency of cropland mapping in areas with complex land cover patterns and agriculture systems.

However, the results also indicate that our model is less accurate in mountainous, urban, and wetland regions, which tend to have higher numbers of false positives. Factors such as slope and shade, urban blocks, and wetland vegetation texture can lead to incorrect predictions. To improve the accuracy of our model, this research suggests several possibilities for future research. Firstly, collecting better-balanced labels among different classes, including increasing the number of non-field labels in

landscapes that result in false positives. Secondly, incorporating slope or DEM data as additional channels to train the model could improve its performance. Thirdly, combining field interior class prediction with field boundary class prediction during accuracy assessment could reduce false negatives. Currently, we only used the field interior class to produce the cropland map. However, as we saw, false negatives of the field class are near the field edge, so combining field interior class prediction with field boundary class prediction could improve accuracy.

Conclusions

This research aims to develop annual cropland maps in dynamic cropland systems at regional to national scales. The use of Google Earth Engine (GEE) enabled the creation of high-resolution, country-wide Planet-Scope NICFI Basemaps with growth-season, off-season, and false color composites to digitize labels. 916 Zambia's crop and non-crop labels created by this research contribute to future mapping efforts. The results demonstrate the value of applying the U-Net model with high-resolution imagery for cropland mapping. It effectively predicts field extent in single and mixed-agriculture types regions, including smallholder fields, large commercial farmland, and slash-and-burn agriculture.

The resulting maps can aid in decision-making processes related to agricultural development and food security by providing a better understanding of the distribution and extent of croplands in Zambia. However, it is important to note that the U-Net cropland map covers only a partial area of the country due to the limitation of research time, and further research is needed to obtain a comprehensive view of Zambia's croplands.

The study suggests future opportunities to improve the model by balancing the number of labels among different classes, incorporating slope or DEM data as additional channels to train the model, combining field interior class prediction with field boundary class prediction to reduce false negatives, and applying different CNN architectures in Zambia, such as DenseNet or ResNet. In conclusion, this research demonstrates the value of using the U-Net model with high-resolution imagery for cropland mapping and provides insights for improving future mapping efforts in Zambia and other sub-Saharan African countries.

Reference

[1] van Ittersum, M.K.; van Bussel, L.G.J.; Wolf, J.; Grassini, P.; van Wart, J.; Guilpart, N.; Claessens, L.; de Groot, H.; Wiebe, K.; Mason-D'Croz, D.; et al. Can Sub-Saharan Africa Feed Itself? Proc. Natl. Acad. Sci. USA 2016, 113, 14964–14969.

[2] Stocking, M. Farming and Environmental Degradation in Zambia: The Human Dimension. Appl. Geogr. 1983, 3, 63–77.

[3] Searchinger, T.D.; Estes, L.; Thornton, P.K.; Beringer, T.; Notenbaert, A.; Rubenstein, D.; Heimlich, R.; Licker, R.; Herrero, M. High Carbon and Biodiversity Costs from Converting Africa's Wet Savannahs to Cropland. Nat. Clim. Chang. 2015, 5, 481–486.

[4] Mutanga, O.; Dube, T.; Galal, O. Remote Sensing of Crop Health for Food Security in Africa: Potentials and Constraints. Remote Sens. Appl. Soc. Environ. 2017, 8, 231–239.

[5] Sitian, X., et al. (2022). Probabilistic Tracking of Annual Cropland Changes over Large, Complex Agricultural Landscapes Using Google Earth Engine, Remote Sensing, 14, 4896.

[6] Azzari, G.; Lobell, D.B. Landsat-Based Classification in the Cloud: An Opportunity for a Paradigm Shift in Land Cover Monitoring. Remote Sens. Environ. 2017, 202, 64–74.

[7] Sweeney, S.; Ruseva, T.; Estes, L.; Evans, T. Mapping Cropland in Smallholder-Dominated Savannas: Integrating Remote Sensing Techniques and Probabilistic Modeling. Remote Sens. 2015, 7, 15295–15317.

[8] Feng, Y.; Zeng, Z.; Searchinger, T.D.; Ziegler, A.D.; Wu, J.; Wang, D.; He, X.; Elsen, P.R.; Ciais, P.; Xu, R.; et al. Doubling of Annual Forest Carbon Loss over the Tropics during the Early Twenty-First Century. Nat. Sustain. 2022, 5, 444–451.

[9] Fritz, S.; See, L.; McCallum, I.; You, L.; Bun, A.; Moltchanova, E.; Duerauer, M.; Albrecht, F.; Schill, C.; Perger, C.; et al. Mapping Global Cropland and Field Size. Glob. Chang. Biol. 2015, 21, 1980–1992.

[10] Estes, L.D.; Ye, S.; Song, L.; Luo, B.; Eastman, J.R.; Meng, Z.; Zhang, Q.;

McRitchie, D.; Debats, S.R.; Muhando, J.; et al. High Resolution, Annual Maps of Field Boundaries for Smallholder-Dominated Croplands at National Scales. Front. Artif. Intell. 2022, 4, 744863.

[11] De Marinis, P.; De Petris, S.; Sarvia, F.; Manfron, G.; Momo, E.J.; Orusa, T.; Corvino, G.; Sali, G.; Borgogno, E.M. Supporting Pro-Poor Reforms of Agricultural Systems in Eastern DRC (Africa) with Remotely Sensed Data: A Possible Contribution of Spatial Entropy to Interpret Land Management Practices. Land 2021, 10, 1368.

[12] Saah, D.; Johnson, G.; Ashmall, B.; Tondapu, G.; Tenneson, K.; Patterson, M.; Poortinga, A.; Markert, K.; Quyen, N.H.; San Aung, K.; et al. Collect Earth: An Online Tool for Systematic Reference Data Collection in Land Cover and Use Applications. Environ. Model. Softw. 2019, 118, 166–171.

[13] Estes, L.D.; Ye, S.; Song, L.; Luo, B.; Eastman, J.R.; Meng, Z.; Zhang, Q.; McRitchie, D.; Debats, S.R.; Muhando, J.; et al. High Resolution, Annual Maps of Field Boundaries for Smallholder-Dominated Croplands at National Scales. Front. Artif. Intell. 2022, 4, 744863.

[14] Fritz, S.; See, L.; Rembold, F. Comparison of Global and Regional Land Cover Maps with Statistical Information for the Agricultural Domain in Africa. Int. J. Remote Sens. 2010, 31, 2237–2256.

[15] Fritz, S.; See, L.; McCallum, I.; Schill, C.; Obersteiner, M.; Van der Velde, M.; Boettcher, H.; Havlík, P.; Achard, F. Highlighting Continued Uncertainty in Global Land Cover Maps for the User Community. Environ. Res. Lett. 2011, 6, 044005.

[16] Jain, M.; Mondal, P.; DeFries, R.S.; Small, C.; Galford, G.L. Mapping Cropping Intensity of Smallholder Farms: A Comparison of Methods Using Multiple Sensors. Remote Sens. Environ. 2013, 134, 210–223.

[17] Lobell, D.B. The Use of Satellite Data for Crop Yield Gap Analysis. Field Crops Res. 2013, 143, 56–64.

[18] Burke, M.; Driscoll, A.; Lobell, D.B.; Ermon, S. Using Satellite Imagery to Understand and Promote Sustainable Development. Science 2021, 371, eabe8628.

[19] Bullock, E.L.; Healey, S.P.; Yang, Z.; Oduor, P.; Gorelick, N.; Omondi, S.;

Ouko, E.; Cohen, W.B. Three Decades of Land Cover Change in East Africa. Land 2021, 10, 150.

[20] Sarvia, F.; De Petris, S.; Borgogno-Mondino, E. Mapping Ecological Focus Areas within the EU CAP Controls Framework by Copernicus Sentinel-2 Data. Agronomy 2022, 12, 406.

[21] Azzari, G.; Jain, M.; Lobell, D.B. Towards Fine Resolution Global Maps of Crop Yields: Testing Multiple Methods and Satellites in Three Countries. Remote Sens. Environ. 2017, 202, 129–141.

[22] Searchinger, T., Waite, R., Hanson, C., Ranganathan, J., Dumas, P., Matthews, E., et al. (2019). Creating a Sustainable Food Future: A Menu of Solutions to Feed Nearly 10 Billion People by 2050. Final Report. Washington, D.C., United States: WRI.

[23] Lobell, D. B., Cassman, K. G., and Field, C. B. (2009). Crop Yield Gaps: Their Importance, Magnitudes, and Causes. Annu. Rev. Environ. Resour. 34, 179–204.

[24] Licker, R., Johnston, M., Foley, J. A., Barford, C., Kucharik, C. J., Monfreda, C., et al. (2010). Mind the gap: How Do Climate and Agricultural Management Explain the 'yield gap' of Croplands Around the World? Glob. Ecol. Biogeogr. 19, 769–782.

[25] Mueller, N. D., Gerber, J. S., Johnston, M., Ray, D. K., Ramankutty, N., and Foley, J. A. (2012). Closing Yield Gaps through Nutrient and Water Management. Nature 490, 254–257.

[26] Estes, L. D., Searchinger, T., Spiegel, M., Tian, D., Sichinga, S., Mwale, M., et al. (2016b). Reconciling Agriculture, Carbon and Biodiversity in a savannah Transformation Frontier. Phil. Trans. R. Soc. B 371, 20150316.

[27] Fritz, S., See, L., McCallum, I., You, L., Bun, A., Moltchanova, E., et al. (2015). Mapping Global Cropland and Field Size. Glob. Change Biol. 21, 1980–1992.

[28] Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., et al. (2012). Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. Remote Sensing Environ. 120, 25–36.

[29] Lesiv, M., Fritz, S., McCallum, I., Tsendbazar, N., Herold, M., Pekel, J.-F., et al.

(2017). November. Evaluation of ESA CCI Prototype Land Cover Map at 20m. Monograph.

[30] Xiong, J., Thenkabail, P., Tilton, J., Gumma, M., Teluguntla, P., Oliphant, A., et al. (2017). Nominal 30-m Cropland Extent Map of Continental Africa by Integrating Pixel-Based and Object-Based Algorithms Using Sentinel-2 and Landsat-8 Data on Google Earth Engine. Remote Sensing 9, 1065.

[31] Zhang, X., Liu, L., Chen, X., Gao, Y., Xie, S., and Mi, J. (2021). GLC_FCS30: Global Land-Cover Product with fine Classification System at 30 M Using Time-Series Landsat Imagery. Earth Syst. Sci. Data 13, 2753–2776.

[32] ESA (n.d.). ESA CCI LAND COVER S2 Prototype Land Cover 20m Map of Africa 2016. Available at: http://2016africalandcover20m.esrin.esa.int/.

[33] Maxwell, A. E., Warner, T. A., and Fang, F. (2018). Implementation of Machine Learning Classification in Remote Sensing: An Applied Review. Int. J. Remote Sensing 39, 2784–2817.

[34] Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., and Johnson, B. A. (2019). Deep Learning in Remote Sensing Applications: A Meta-Analysis and Review. ISPRS J. Photogrammetry Remote Sensing 152, 166–177.

[35] Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., et al. (2015). Global Land Cover Mapping at 30m Resolution: A POK-Based Operational Approach. ISPRS J. Photogrammetry Remote Sensing 103, 7–27.

[36] Stehman, S. V., and Foody, G. M. (2019). Key Issues in Rigorous Accuracy Assessment of Land Cover Products. Remote Sensing Environ. 231, 111199.

[37] Van Vliet, N., Mertz, O., Birch-Thomsen, T., and Schmook, B. (2013). Is There a Continuing Rationale for Swidden Cultivation in the 21st Century? Hum. Ecol. 41, 1–5.

[38] CIAT, World Bank (2017) Climate-Smart Agriculture in Zambia. CSA Country Profiles for Africa Series. International Center for Tropical Agriculture (CIAT), Washington, D.C. 25 p.

[39] Indaba Agricultural Policy Research Institute (2020) *Zambia Agriculture Status*

Report 2020.

[40] Chikowo, R. Description of cropping systems, climate, and soils in Zambia, *Global Yield Gap Atlas*. Available online: https://www.yieldgap.org/zambia(accessed on 27 May 2022).

[41] Ellenbroek, G.A., (1987) Ecology and productivity of an african wetland system, Springer Science & Business Media.

[42] Lesiv, M. el at. (2019) Estimating the global distribution of field size using crowdsourcing. Global Change Biology, 25, 174-186.

[43] Chen, J., Ban, Y., and Li, S. (2014). China: Open access to Earth Landcover Map, *Nature*, 514(7523), 434–434.

[44] ESA. A cropland layer derived from the ESA CCI LC map at a 300 m resolution for 2015. Available online: https://www.esa-landcover-cci.org/ (accessed on 19 May 2022).

[45] Waldner, F. et al. (2016) A Unified cropland layer at 250 m for global agriculture monitoring, *Data* (1), 3.

[46] Fritz, S. et al. (2015) Mapping global cropland and field size, *Global Change Biology,* 21(5): 1980–1992.

[47] Estes, L. et al. (2022) Creating Open Agricultural Maps & Groundtruth Data to Better Deliver Farm Extension Services February Report.

[48] Ronneberger, O., P. Fischer, and T. Brox. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation, in N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Springer International Publishing, 234–241, Cham.

[49] Kattenborn, T., Leitloff , J., Schiefer, F. and Hinz, S. (2021) Review on Convolutional Neural Networks (CNN) in vegetation remote sensing, *ISPRS Journal of Photogrammetry and Remote Sensing*, 173: 24-49.

[50] Ronneberger, O., Fischer, P., Brox, T. (2021) U-Net: Convolutional Networks for Biomedical Image Segmentation, MICCAI 2015: Medical Image Computing and Computer-Assisted Intervention, 234–241.

[51] Abdi, O., Uusitalo, J. and Kivinen, V. (2022) Logging Trail Segmentation via a Novel U-Net Convolutional Neural Network and High-Density Laser Scanning Data, *Remote Sensing*, 14(2): 349.

[52] Chollet, F. (2018) Deep Learning with Python; Manning Publications Co, Shelter Island, NY, USA, ISBN 1617294438.

[53] Neubert, P., and Protzel, P. (2014) Compact Watershed and Preemptive SLIC: On Improving Trade-Offs of Superpixel Segmentation Algorithms, in *2014 22nd International Conference on Pattern Recognition*, Stockholm, Sweden: IEEE), 996–1001.

[54] Visvalingam, M., and Whyatt, J. D. (1993) Line Generalisation by Repeated Elimination of Points, *Cartographic J* (30): 46–51.

[55] Stehman, S. V., and Foody, G. M. (2019) Key Issues in Rigorous Accuracy Assessment of Land Cover Products, *Remote Sensing Environment* (231), 111199.

[56] Olofsson, P., Foody, G. M., Stehman, S. V., and Woodcock, C. E. (2013) Making Better Use of Accuracy Data in Land Change Studies: Estimating Accuracy and Area and Quantifying Uncertainty Using Stratified Estimation, *Remote Sensing Environ* (129): 122–131.

[57] Olofsson, P., Foody, G. M., Herold, M., Stehman, S. V., Woodcock, C. E., and Wulder, M. A. (2014) Good Practices for Estimating Area and Assessing Accuracy of Land Change, *Remote Sensing Environ* (148): 42–57.

[58] Stehman, S. V., and Foody, G. M. (2019) Key Issues in Rigorous Accuracy Assessment of Land Cover Products, *Remote Sensing Environ* (231), 111199.

[59] Aston, Z, Zachary, C. L., Mu, L., Alexander J. S. (2022) Dive into Deep Learning.

[60] Story M, Congalton G R (1986) Accuracy assessment: a user's perspective [J]. Photogrammetric Engineering and Remote Sensing, 52: 397-399

[61] Gal, Y.; Ghahramani, Z. (2016) Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. In Proceedings of the 2016, ICM'16, New York, NY, USA, 19–24 June 2016; pp. 1050–1059.

[62] Nesterov, Y. (2018). Lectures on convex optimization. Vol. 137. Springer.

[63] Estes, L., Chen, P., Debats, S., Evans, T., Ferreira, S., Kuemmerle, T., et al. (2018). A Large-area, Spatially Continuous Assessment of Land Cover Map Error and its Impact on Downstream Analyses. Glob. Change Biol. 24, 322–337. doi:10.1111/gcb.13904

[64] Zambia Wildlife Authority (2009) Zambezi Floodplain Ramsar Information Sheet" Archived March 26, 2009, Wayback Machine, in Ramsar Sites Information Service, Wetlands International. https://web.archive.org/web/20090326150707/http://www.wetlands.org/reports/ris/1ZM007_2007.pdf

[65] Malambo, F. M., Syampungani, S., (2008) Opportunities and challenges for sustainable management of miombo woodlands: the Zambian perspective, Working Papers of the Finnish Forest Research Institute 98: 125–130.

[66] LeCun, Y., Bengio, Y., & others. (1995). Convolutional networks for images, speech, and time series. The handbook of brain theory and neural networks, 3361(10), 1995.

[67] Nair, V., & Hinton, G. E. (2010). Rectified linear units improve restricted boltzmann machines. Icml.

[68] Dumoulin, V., & Visin, F. (2016). A guide to convolution arithmetic for deep learning. arXiv preprint arXiv:1603.07285.

[69] Jordan, J. (2018) An overview of semantic image segmentation, https://www.jeremyjordan.me/semantic-segmentation/.

[70] Python Documentation (2023) pickle — Python object serialization, https://docs.python.org/3/library/pickle.html.

[71] Wang, Z., O'Boyle, M. (2018) Machine Learning in Compiler Optimisation, IEEE.

[72] P. Potapov, S. Turubanova, M.C. Hansen, A. Tyukavina, V. Zalles, A. Khan, X.-P. Song, A. Pickens, Q. Shen, J. Cortez. (2021) Global maps of cropland extent and change show accelerated cropland expansion in the twenty-first century. Nature

Food. https://doi.org/10.1038/s43016-021-00429-z

[73] Azzari, G., & Lobell, D. B. (2017). Landsat-based classification in the cloud: An opportunity for a paradigm shift in land cover monitoring. Remote Sensing of Environment, 202, 64-74.

[74] Xiong, J., Thenkabail, P., Gumma, M., Teluguntla, P., Poehnelt, J., Congalton, R., Yadav, K., and Thau, D. (2017) Automated cropland mapping of continental Africa using Google Earth Engine cloud computing, ISPRS Journal of Photogrammetry and Remote Sensing, 126: 225-244.

[75] Jayne, T. S., Chamberlin, J., and Headey, D. (2014) Food Policy, 48: 1-17.

[76] Karamage, F. et al. (2016) Extent of Cropland and Related Soil Erosion Risk in Rwanda, Sustainability, 8(7) 609.

[77] Tao, B. et al. (2013) Terrestrial carbon balance in tropical Asia: Contribution from cropland expansion and land management, Global and Planetary Change, 100: 85-98.

[78] See, L. et al. (2015) Improved global cropland data as an essential ingredient for food security, Global Food Security, 4: 37-45.

[79] Ngoma, H., et al. (2021) Climate-smart agriculture, cropland expansion and deforestation in Zambia: Linkages, processes and drivers, 107.

[80] United Nations Conference on Trade and Development (UNCTAD). An Investment Guide to Zambia: Opportunities and Conditions 2011; United Nations: Geneva, Switzerland, 2011.

[81] Wang, S.; Chen,W.; Xie, S.M.; Azzari, G.; Lobell, D.B.Weakly Supervised Deep Learning for Segmentation of Remote Sensing Imagery. Remote Sens. 2020, 12, 207.

[82] Waldner, F., and Diakogiannis, F. I. (2020) Deep learning on edge: Extracting field boundaries from satellite images with a convolutional neural network, Remote Sensing of Environment, 245.

[83] Korstanje, J., (2021) The F1 score, https://towardsdatascience.com/the-f1-score-bec2bbc38aa6

[84] Boston, T., et al. (2022) Comparing CNNs and Random Forests for Landsat

Image Segmentation Trained on a Large Proxy Land Cover Dataset, Remote Sensing, 14.

[85] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. The Journal of Machine Learning Research, 15(1), 1929–1958.

[86] Oleszak, M., (2020) Monte Carlo Dropout, Medium, https://towardsdatascience.com/monte-carlo-dropout-7fd52f8b6571.

[87] Gal Y. & Ghahramani Z., (2016) Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning, Proceedings of the 33rd International Conference on Machine Learning.

[88] Pandey, P., et al. (2023) PLANET BASEMAPS FOR NICFI DATA PROGRAM ADDENDUM TO BASEMAPS PRODUCT SPECIFICATION.

[89] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classifcation with deep convolutional neural networks. Adv Neural Inf Process Syst. 2012;25:1106–14.

[90] Shorten, C., Khoshgoftaar, T.M., 2019. A survey on image data augmentation for deep learning. J. Big Data 6, 1–48.

[91] Song, L., Estes, A. B., Estes, L. D. (2023) A super-ensemble approach to map land cover types with high resolution over data-sparse African savanna landscapes, International Journal of Applied Earth Observations and Geoinformation, 116.

[92] Simonyan, K., Zisserman, A. (2015) VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION, Computer Vision and Pattern Recognition.