MAPPING SOYBEAN EXPANSION AND THE IMPACT ON FOOD SECURITY AMONG SMALLHOLDER FARMERS IN ZAMBIA

BRIGHT CHOTA NKOLE

Enschede, The Netherlands, JULY, 2023

SUPERVISORS:

Dr. Yue Dou

Dr. Mariana, Belgiu

MAPPING SOYBEAN EXPANSION AND THE IMPACT ON FOOD SECURITY AMONG SMALLHOLDER FARMERS IN ZAMBIA

BRIGHT CHOTA NKOLE

Enschede, The Netherlands, JULY, 2023

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.

Specialization: Natural Resources Management

SUPERVISORS:

Dr.Yue Dou

Dr. Mariana Belgiu

THESIS ASSESSMENT BOARD:

Dr Ir.Anton Vrieling (Chair)

Dr Zhanli Sun, (External Examiner, Leibniz Institute of Agricultural Development in Transition Economies)

Disclaimer

Abstract

The rapid expansion of soybean cultivation and the impact on food security in southern Africa remain inadequately documented and understood. Existing studies predominantly concentrate on assessing the income generation potential at the household level, neglecting the broader implications for food security. This study utilized remote sensing (RS) techniques and Geographic Information Systems (GIS) to analyze the expansion of soybean farming in the Chibombo district of Zambia. High-resolution satellites of Sentinel-1 and Sentinel-2 data along with the ground-based data were integrated into random forest machine learning algorithms to accurately map different land use/land cover types for the years 2017, 2020, and 2023 respectively providing reliable information on land use/land cover changes. The overall accuracy for the 2023 classification results was achieved at 86% indicating a significant expansion of soybean cultivation by 156.6% from 2017 to 2023 while other food crops showed contrasting trends. The classifier model was transferred to the years 2017 and 2020, for which ground-based reference data was unavailable. The classification results were validated using statistical data from the crop focus survey (CFS) with no significant differences in cultivated soybean areas. This approach enabled the study to accurately identify soybean cropland areas and assess the implications of soybean expansion on food security at the household level.

To assess the impact of soybean expansions on the food security situation in the study area, two scores, namely the Household Dietary Diversity Score (HDDs) and the Household Food Insecurity Access Score (HFIAs) were used to measure food security at the household level. The results for HDDs revealed that a significant proportion of households in the Chibombo district have low or moderate dietary diversity. 40% of households had a low dietary diversity score, 40.4% had a moderate score, and only 19.6% had a high score. The HFIAs indicated a high prevalence of food insecurity among smallholder farmers in the study area. 50.8% of households were severely food insecure, 28.7% were food secure, and 20.4% were moderately food insecure. This suggests that a substantial number of households are facing challenges in accessing sufficient and nutritious food. The spatial distribution analysis of HDDs and HFIAs in each ward, considering soybean cultivation areas, showed discrepancies in dietary diversity and food accessibility. While some wards exhibited a higher percentage of households with low dietary diversity, others demonstrated greater diversity. Similarly, food insecurity levels varied across wards, with some facing substantial challenges while others had relatively higher food security. Interestingly, the study found no clear correlation between soybean cultivation area and HDDs or HFIAs. Wards with larger soybean areas did not consistently show better food security outcomes, suggesting a potential mismatch between soybean expansion and household food security indicators. Additionally, significant disparities in dietary diversity and food accessibility were identified among the various categories of major stakeholders involved in soybean production namely: the households who engaged in growing, expansions, selling, and utilizing the soybean commodity after harvesting.

Across all categories, common elements such as the interquartile range, median values, and potential disparities in both HDDs and HFIA scores provided valuable insights into the distribution, central tendency, and variations in dietary diversity and food insecurity within each category. The results also indicated that there was no clear trend indicating that wards with larger soybean areas consistently had better HDDs and HFIAs. The overall findings of this study indicate that the expansion of soybean cultivation has a limited impact on enhancing dietary diversity and improving food security among smallholder farmers. This study provides a valuable foundation for informing nutrition-sensitive agriculture policies, about the effects of soybean agricultural expansions on the food security of smallholder farmers.

Acknowledgements

I would like to express my sincere gratitude to my supervisors, Dr Y. Dou and Dr M. Belgiu, for their invaluable guidance, consistent feedback, and words of encouragement throughout my research journey. Their support has been invaluable in helping me accomplish this study despite challenging circumstances. Additionally, I would like to extend my deep appreciation to Dr Raymod G. Nijmeijer, the NRM Program Director, for his caring and attentive approach during my two years of master's studies, particularly during the thesis phase. Furthermore, I am thankful to Dr Ir. Anton Vrieling, the Chairman of this research work, for his valuable insights and constructive feedback.

I would like to extend my heartfelt gratitude to the Faculty of Geo-information Science and Earth Observation (ITC) at the University of Twente and the Orange Knowledge Programme (OKP) for their generous financial support. Their assistance has been instrumental in enabling me to pursue my MSc degree and undertake this research. Their contributions have played a significant role in the successful completion of this study.

Moreover, I would like to express my deepest gratitude to my family and friends in Zambia for their unwavering support throughout my journey, despite the physical distance between us. Their constant encouragement and understanding have been invaluable to me, and I am truly grateful for their presence in my life. Additionally, I am indebted to the field officers who generously assisted me during the fieldwork phase, enabling the collection of vital ground truth data. Their contributions have greatly enhanced the quality and reliability of this study.

Lastly, I want to convey my deepest appreciation to my beloved partner, Euphemia Chama, my daughter Monica Chota, and my son Brown Chota for their unwavering support throughout this adventurous journey. Their presence, ability to bring laughter amidst the chaos, and constant reminder of perspective have been my anchor throughout this journey, and their support has meant the world to me. Amid challenging circumstances and demanding research work, their unwavering love and positive energy have provided me with a sense of stability and encouragement. I am deeply grateful for their unwavering presence and the profound impact they have had on my life and this adventure.

Table of contents

Acronyms

CFS	Crop Focus Survey
GEE	Google Earth Engine
GIS	Geographic Information Systems
HDDs	Household Dietary Diversity Score
HFIAs	Household Food Insecurity Access Score
NDRE	Normalized Difference Vegetation Index with the red edge
NDVI	Normalized Vegetation Index
NDWI	Normalized Difference Water Index
NIR	Near Infrared
RF	Random forest
VH	Vertical-Horizontal Polarization
VV	Vertical-Vertical Polarization
OA	Overall accuracy
PA	Producer accuracy
SAR	Significant research gap in the Southern African Region
UA	User accuracy

List of Figures

List of tables

1 Introduction

1.1 Background

The increasing demand for soybean as a source of protein has led to a significant expansion of its production worldwide (Voora et al., 2020). Globally, the soybean cropland area has rapidly expanded over the last decades, from 74 million hectares (ha) in 2000 to 129 million ha in 2021. Extensive studies have been conducted on soybean production and expansion, primarily in South America, North America, and Asia (Pacheco, 2012). However, there is a notable research gap in the Southern African Region (SAR) regarding the expansion of soybean cultivation and its potential impact on food security in the region (Siamabele, 2021). This gap hinders understanding the region's socio-economic factors that influence soybean expansion. As a result, the potential benefits of sustainable soybean expansion, such as improved food security, remain largely untapped in SAR (Khojely et al., 2018).

The rapid expansion of soybean cultivation has raised concerns regarding its potential impact on food security, particularly among smallholder farmers in the SAR. (Siamabele, 2021). In countries such as South Africa, Zambia, and Zimbabwe, the cultivation of soybean has increased dramatically due to the growing need for protein-rich food and feed (Siamabele, 2019). The area under cultivation in SAR expanded from 134,150 in 2000 to 827,100 ha in 2021 (Figure *1*). South Africa emerged as the largest producer, recording the majority of the expansions, followed by Zambia. (Khojely et al., 2018). The expansion has also resulted in the conversion of fertile lands previously used for growing food crops to soybean fields, raising worries about the region's future food security (Savala et al., 2022). Therefore, understanding the expansion of cash crops such as soybean cultivation in the SAR is important for several reasons, especially concerning food security threats (Frelat et al., 2016). Some of the reasons include competition with food crops that can lead to the displacement of food crops, including staple crops such as maize and cassava, which can contribute to food insecurity in the region by reducing the availability of locally grown food (Frelat et al., 2016). The other reason is that most soybean producers in the SAR region are smallholder farmers, particularly vulnerable to food insecurity (Khojely et al., 2018). Understanding the factors driving soybean expansion is crucial for ensuring food security and protecting smallholder farmers' livelihoods.

Remote sensing (RS) and Geographic Information Systems (GIS) provide powerful tools for understanding the factors that drive soybean expansion among smallholder farmers in the SAR through mapping the expansion. This allows for a better understanding of the drivers behind soybean expansion in the region. (Xu et al., 2018). Remote sensing data such as Sentinel-1(S1) and Sentinel-2 (S2) can provide high-resolution images of the earth's surface that can be used to map and monitor changes in land use or cropland (Rao et al., 2021). The temporal resolution of remote sensing data is instrumental in monitoring temporal changes over time (Gikov et al., 2019), additionally, it can provide important information on the impact of soybean expansion on food security in the region.

However, mapping crop types like soybean using remote sensing in the SAR poses significant challenges. The region's high cloud cover obstructs optical satellite imagery, limiting access to cloud-free data necessary for accurate mapping (Bégué et al., 2020). The diverse landscapes, characterized by smallholder farms and mixed cropping systems, make it more difficult to differentiate crop fields from other land cover types. Moreover, the scattered nature of small-scale crop fields with irregular shapes intermingled with other land types further complicates the mapping process (Bégué et al., 2020). For example, soybean's phenological variability, influenced by varying maturity durations and management practices, adds complexities to its accurate detection and mapping. Additionally, interpreting and validating remote sensing data heavily relies on ground-based observations and field surveys, which may be constrained in certain areas (Bégué et al., 2020). Compared to well-established systems in South America, North America, and Asia with robust RS and GIS capabilities, extensive data availability, and advanced modelling techniques (Wang et al., 2020), the level of accuracy in mapping soybean expansion lags in many counties of the SAR. This disparity stems from insufficient ground-based validation data and less developed mapping frameworks within the region (Bégué et al., 2020). Consequently, the insufficient knowledge regarding the spatial patterns and dynamics of soybean expansion in the SAR hinders the effective management of sustainable cropland expansion and its potential contribution to food security.

In recent years machine learning algorithm such as random forest (RF) has proven to be robust and accurate approach for classifying cropland (You & Dong, 2020) and assessing the impact of cropland expansion on food security among smallholder farmers in SAR (Phalan et al., 2013). Characteristics such as robustness to outliers, missing data, and unbalanced datasets outperform many other classifiers including maximum likelihood (You & Dong, 2020). Through the integration of RS and machine learning, researchers and policymakers can gain new insights into the distribution of unsustainable cropland expansion and the impact on food security in the region (Li et al., 2021). Furthermore, researchers and policymakers can effectively link spatial data on cropland expansion and land use change to data on food security and household livelihoods by combining demographic and socioeconomic factors (Li et al., 2021). These factors encompass understanding population characteristics and assessing poverty levels, which can be captured through household food security surveys (Mango et al., 2014).

This study, therefore, aims to map the expansion of soybean cultivation in Zambia located in SAR and evaluate its impact on food security at the household level among smallholder farmers. Mapping soybean expansions in Zambia is an important study due to the country's rapid soybean expansion and increasing production and its potential impacts on food security. From 2001 to 2021, the cultivation area of soybeans in Zambia witnessed a remarkable surge, expanding from 3,889 to 311,254 ha (Figure *1*). This substantial increase highlights the rapid expansion of soybean cultivation in Zambia over the past few decades, while the production also showed an increasing trend from 2,350 in 2001 to 411,115 metric tons in 2021. This indicates a growth in soybean production, corresponding to the expansion of the cultivation area (Figure 1). Therefore, understanding the spatial patterns and dynamics of soybean expansion is crucial for assessing its contributions to food security and protecting the livelihoods of smallholder farmers.

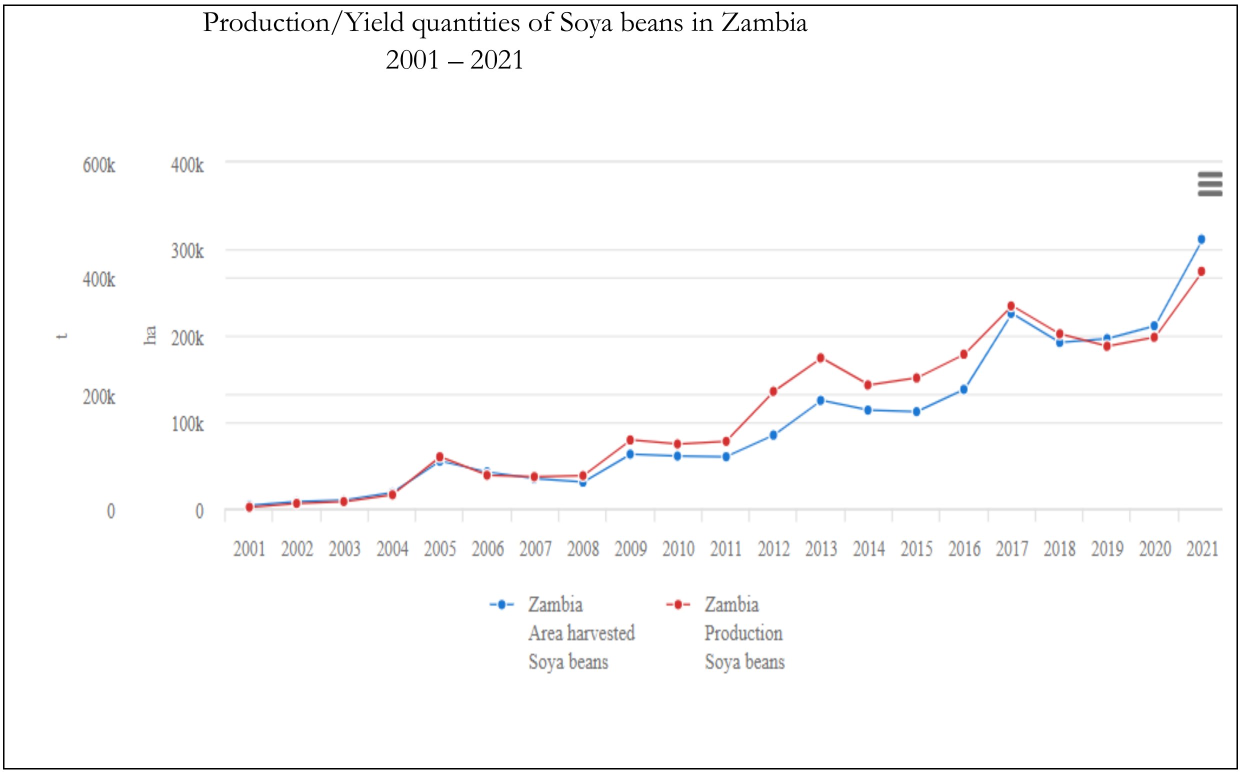


Figure 1. Soybean area harvested (ha) and production (tonnes) in Zambia from 2001 to 2021 (FAOSTAT June 06, 2023).

Mapping soybean expansion in Zambia will provide a significant understanding of the spatial distribution and extent of soybean cultivation, supporting policymakers in monitoring the expansion and making well-informed decisions. By gaining valuable knowledge about the patterns and trends associated with soybean expansion, targeted interventions and strategies can be developed to maximize the positive impacts while mitigating any potential negative consequences on food security among smallholder farmers.

1.2 Objectives

This study aims at mapping the expansion of soybean cultivation and assesses its impact on food security among smallholder farmers in Zambia (using Chibombo District, Zambia as a case study) by answering the following research questions:

(a) What is the extent of soybean expansion in the Chibombo district and how has it changed over time?
(b) How has soybean expansion affected the dietary diversity of smallholder farmers?
(c) How has the expansion of soybean affected the food insecurity among smallholder farmers?

2 Data and methods

2.1 Study area

The study area is located in the central province of Zambia (Figure *2*), Chibombo district with a total area of 13,423 km^2 (Latitude: 14° 39' 19.5408" Longitude: 28° 5' 19.8888" E). The district benefits from a favourable climate, characterized by moderate and evenly distributed rainfall ranging from 800 to 1000 mm annually. Additionally, the altitude of the area ranges from 1100 to 1200 meters above sea level. (Mubanga & Bwalya Umar, 2020a). It experiences three distinct seasons: the warm and wet rainy season from December to April, the cool and dry season from May to August, and the hot and dry season from September to November. The topography of the area is characterized by gently undulating terrain. The soil conditions in the area are characterized by fertile sandy-loam soils, known for their optimal composition of sand, silt, and clay. (Chilambwe et al., 2022).

The soils possess high fertility and effective moisture retention capabilities, making them suitable for agricultural production (Chilambwe et al., 2022). The combination of both favourable climatic and soil conditions makes the Chibombo district a well-suited location for the agricultural production of common legumes and cereals. The current population of the Chibombo district is 421,315, of which 50.5% are males and 49 % are females (Chilambwe et al., 2022). Within this total population, an estimated 48,000 to 55,000 smallholder farmers are distributed across the 20 agricultural camps found in 16 wards of the district (Mubanga & Bwalya Umar, 2020b). The district depends on smallholder farmers for agricultural production, with maize, soybean, cotton, groundnuts, and sunflower being the key crops cultivated. (Mubanga & Bwalya Umar, 2020b).

Chibombo district in Zambia has witnessed a significant increase in soybean cropland, with a remarkable growth rate of approximately 108% from the 2016/2017 to 2021/2022 farming seasons. The cropland for soybean expanded from 14,588 to 30,385 ha (Zam-stats). This expansion rate exceeds the expansion rates observed in other districts across the country, highlighting the rapid expansion and adoption of soybean cultivation in the Chibombo district (Zam-stats). However, the district faces challenges regarding food security, particularly among communities dependent on subsistence agriculture for their livelihoods (Sebatta et al., 2014). Various factors contribute to the food insecurity situation in the district. These include low-income levels, limited crop diversification, and restricted access to essential agricultural inputs like fertilizers and certified seeds (Mainza, 2022). The situation is further compounded by the decline in cropland area for some food crops, such as cassava and sorghum which supplements micronutrients and prevents stunting (Bwalya, 2022).

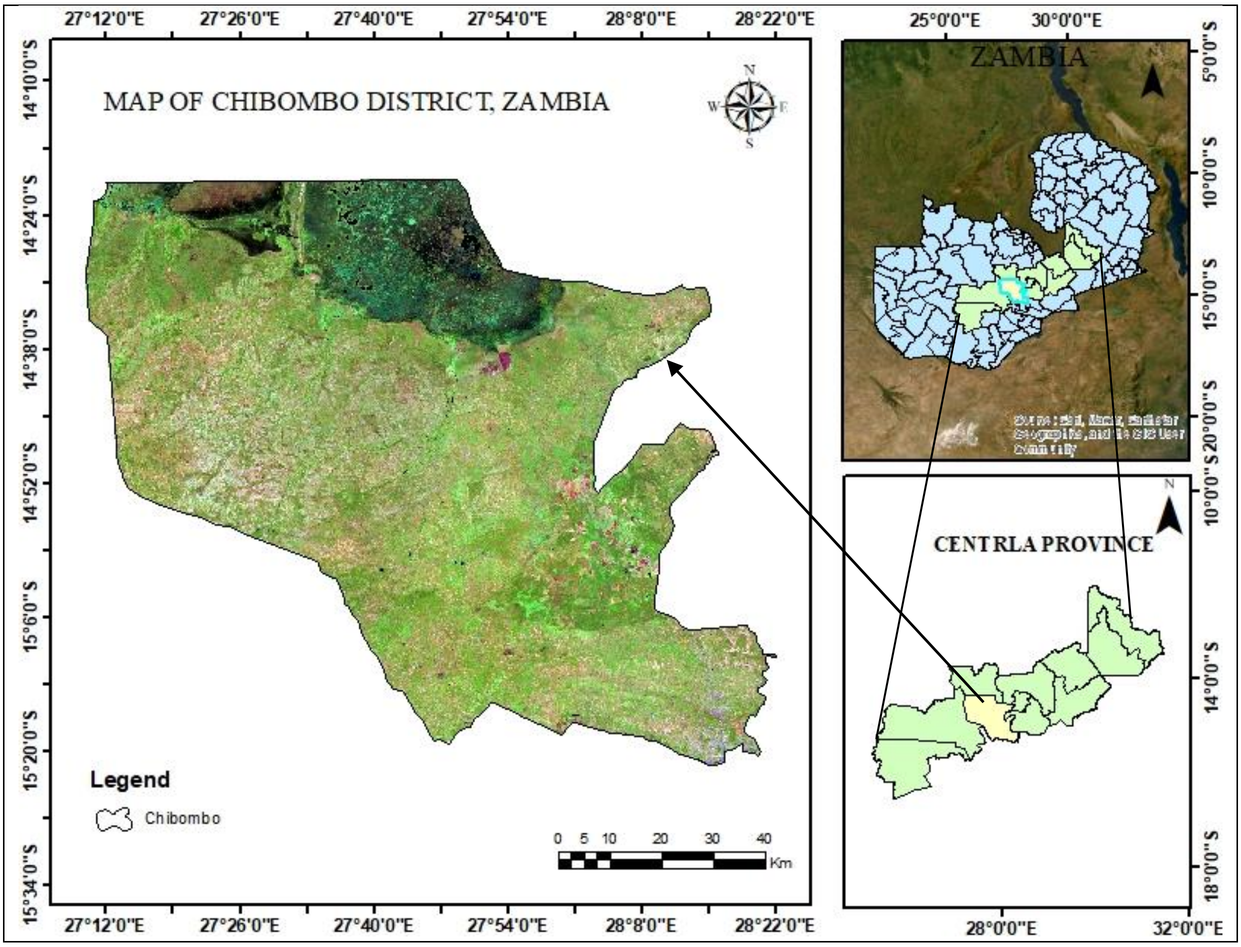


Figure 2. Study area map

2.2 Data and processing

The study utilized satellite images from Sentinel-1 (S1) and Sentinel (S2) obtained through the Google Earth Engine platform (GEE) to track changes in soybean cropland in the study area. Ground reference data in the form of polygons representing major crops, forests, built-up/bare soils, and water bodies were collected for accurate crop type and land cover classification. Additionally, a household food security survey was conducted to assess the food security status of smallholder farmers, collecting data on socio-economic variables and specific information related to soybean cultivation. The combination of satellite imagery, ground reference data, and the food security survey aimed to provide a thorough understanding of the dynamics of soybean expansion, land cover classification, and their impact on household food security in the study area.

2.2.1 Satellite images

S1 and S2, which are part of the Copernicus program, are Earth Observation satellites operated by the European Space Agency (ESA) (Xie & Niculescu, 2022). Both S1 and S2 data images were acquired using GEE, a cloud-based platform that offers access to a vast array of satellite imagery and geospatial information (Liang et al., 2023). To align with the agricultural calendar of the study area, the data was filtered to cover the growing season, which starts in December and ends in August. Image collections were acquired every month from 15th November to 10th June in 2017, 2020 and 2023 to track changes in the soybean cropland over time.

The study area is located in the tropics, a region, characterized by a high frequency of cloud cover significantly limiting the availability of optical imagery (Talema & Hailu, 2020). To mitigate the challenges posed by frequent cloud cover in the study area, the classification process involved the utilization of S1 (Table 1). Compared to S2, an optical sensor, S1 is less affected by cloud cover (Talema & Hailu, 2020) and was able to provide data for the growing season between December to March, thus filling a critical image availability for the study area. The images for S1 were filtered to include only images with both Vertical-Vertical polarization (VV) and Vertical-Horizontal polarization (VH) and in Interferometric Wide (IW) instrument mode (Planque et al., 2021). Images with low-quality areas were filtered out using a threshold value of -20.0 and the median composite of all the images in the selected period was obtained. The resulting images were clipped to the area of interest (aoi) and the index of the ratio of VV and VH bands was derived as an input into the classification (Yu et al., 2021). The ratio for S1 data is typically calculated as the ratio between the backscatter coefficient in the Vertical (VV) polarization and the backscatter coefficient in the Horizontal (VH) polarization (S1 ratio = VV/VH) (Soudani et al., 2021).

Table 1: Summary description of S1 features used in the crop type/ landcover classification in this study in the growing season (from December to March) in 2017, 2020 and 2023.

Acquisition period (Monthly)	Processing level	Number of images	Band and indices
Dec, Jan, Feb & Mar	Ground Range Detected (GRD)	5 (median composite)	VH, VV, VH/VV

The images for the S2 were filtered by selecting images within the period of interest from April, May and June while pre-filtering to obtain images with low cloud cover percentage-Less than 10% (Simonetti et al., 2021). The bands of interest included the blue, green, red, red edge-1, red edge-2, red edge 3, NIR, SWIR1, and SWIR2 (Table 3) To remove the clouds to the desired percentage level, the function "mask2sclouds" was applied to the images (Simonetti et al., 2021). Subsequently, monthly composite images were created for each band, and the median value of each band was computed for all the images (Table 2). These computed values were then stacked together to create a multi-band image, which was further clipped to our study area (Rao et al., 2021). The normalized vegetation index (NDVI) = (NIR - Red) / (NIR + Red) (Tucker, 1979), normalized difference vegetation index with the red edge (NDRE) = (NIR - RE) / (NIR + RE) (Doumit & Kiselevm, 2017), and the normalized difference water index (NDWI) = (Green - NIR) / (Green + NIR) (McFeeters, 1996) was calculated using the multi-band image.

Table 2: Summary description of S2 features used in the crop type/ landcover classification in this study in the growing season (from April to May) in 2017, 2020 and 2023.

Acquisition period (Monthly)	Processing level	Number of images used	Indices used
April, May& June	Level 2A	(3) Monthly composite	NDVI, NDRE, & NDWI

The complementary data of NDVI, NDRE, and NDWI in combination with other spectral bands, the ground truth on crops and landcover were used to help improve the accuracy of the cropland classification model of the RF algorithm and provide a more detailed understanding of the spatial-temporal dynamics (Blickensdörfer et al., 2022).

Table 3: Summary description of the S2 bands used in the classification.

Sentinel-2	Spatial resolution	Central wavelength
Band-2 Blue	10	490 nm
Band-3 Green	10	560 nm
Band-4 Red	10	665 nm
Band-5 Vegetation red edge	20	705 nm
Band-6 Vegetation red edge	20	740 nm
Band-7 Vegetation red edge	20	783 nm
Band-8 Near infra-red	10	842 nm
Band-8A Vegetation red edge	20	865 nm
Band 9 Short wave infrared	60	940 nm
Band 10 Short wave infrared	60	1375 nm
Band 11 Short wave infrared	20	1610 nm
Band 12 Short wave infrared	20	2190 nm

2.2.2 Ground truth data, training samples, and validation samples

Polygons of major crops of soybean, maize, sunflowers, groundnuts, and cotton as well as forests, built-up areas, and water bodies were collected for crop type/landcover classification across the agricultural camps dominated by smallholder farmers (Table 4). The collected polygons served as ground reference data for crop type and land cover classification. The ground truth data was used to train (70%) and validate (30%) the RF classification algorithm for accurate crop type/land cover mapping using satellite imageries (Rao et al., 2021). The ground reference data collection took place from January to March 2023 using a Qfield app, a mobile data collection and management application compatible with QGIS, a widely-used open-source desktop GIS software (J. Duncan et al., 2022).

Table 4: Summary description of the cropland/landcover samples.

Type	Maize	Soybeans	Sunflower	Cotton	Groundnuts
polygons	378	162	32	21	18

2.2.3 Food security survey

Food security is a state where all people must have constant access to a sufficient, secure, and nourishing food supply that satisfies their dietary needs and personal preferences to maintain an active and healthy lifestyle (Awoyemi et al., 2022). Food security encompasses the four dimensions of availability, accessibility, utilization, and stability of food at various levels, including the individual, household, national, and global levels. (Castell et al., 2015). To assess the food security status of smallholder farmers in Zambia, a survey was conducted specifically focusing on household food security. This survey served as a significant tool for analysing the food security situation in the study area. A food security survey is a widely employed research instrument utilized to evaluate the food security situation of individuals, households, or communities. (World Food Program, 2014). It involves collecting data on various aspects related to food availability, access, utilization, and stability, to understand the extent and nature of food insecurity and identify potential interventions (Awoyemi et al., 2023). Therefore, the purpose of the food security survey was to identify the specific factors that contribute to food insecurity, including demographics and socioeconomic factors (Mango et al., 2014). To ensure a random selection of smallholder farmers for the household food security survey, a cluster sampling method was employed by using wards as clusters (Kenefick, 2004). The study area comprised 16 wards, and 15 questionnaires were administered per ward, resulting in a total of 240 questionnaires. This approach aimed to include a representative sample of smallholder farmers engaged in soybean cultivation within the study area. The survey covered a range of traditional demographic and socio-economic variables (Table 9) as determinants of food security at the household level (Usman & Haile, 2022). These important socio-economic variables included the age of the household head, gender, education level of the household head, household size, sources of income and food, assets ownership and livestock, distance to the main road and market, and access to farming inputs. Additionally, specific data related to soybean cultivation, including the number of crops grown per household, expansion of soybean, utilization of harvested soybean, selling of soybean for household income, and cropland expansion for soybean, were collected to assess the contribution of soybean cropland expansion to household food security. The data was captured using the KoboToolBox for data management., KoboToolBox simplifies data management by automatically organizing and storing collected data in a centralized database. It allows for data export in various formats, facilitating data analysis and integration with other analysis tools such as r studio and SPSS (Statistical Package for the Social Sciences).

2.3 Methods

2.3.1 Random Forest

In this study, the selection of the RF was crucial for the accurate mapping of soybean expansion. RF is a machine-learning technique known for its high accuracy and robustness, making it particularly effective in handling complex and large datasets (Breiman, 2001). Unlike other classifiers like traditional decision trees, RF is capable of effectively capturing non-linear relationships and interactions between variables, which are commonly found in crop-type mapping (Wang et al., 2020). Additionally, RF can easily accommodate missing data, noisy data, and unbalanced datasets, common obstacles in accurately mapping crop types (Sun et al., 2020). By utilizing the strengths of RF, this study was able to mitigate the challenges and achieve more reliable and accurate results. The RF classifier operates by creating a collection of decision trees, where each tree is trained on a randomly chosen subset of the data. (Breiman, 2001). The classifier combines the predictions of these individual trees to make a final prediction (Breiman, 2001).

Given the large geographic region covered by the study area, the study opted for the GEE platform, due to its capabilities in handling large-scale analysis and processing of geospatial data (Ghosh et al., 2022). The cloud computing platform of GEE allows users to access, manipulate, and analyze massive volumes of geographical datasets in real-time (Liang et al., 2023). The platform offers a comprehensive collection of public data, including an extensive catalogue of satellite imagery and geographic data from various sensors. This data catalogue, which is regularly updated with new images daily, provides free accessibility to all users. (Liang et al., 2023). GEE is known for its user-friendly interface and has gained popularity within the RS community in recent years (Ghosh et al., 2022). It has provided valuable support to various earth observation studies conducted at local, regional, and global levels. In recent years, the implementation of RF in GEE has become increasingly popular for tasks such as land use monitoring, mapping land cover dynamics, and detecting changes in cropland expansions, irrigated areas, and pasture areas. (Ghosh et al., 2022).

To enhance the generalization of individual decision trees within the RF and mitigate overfitting, certain parameters were set in the GEE implementation (You & Dong, 2020). The minimum size of a terminal node parameter was set at 10, specifying the minimum number of samples required for a terminal node. This helps prevent overfitting and encourages the generalization of individual decision trees. A total of 1000 decision trees were used in the analysis as previous studies by Belgiu & Csillik, (2018) suggest that a larger number of trees generally leads to improved performance and does not result in an increase in the number of errors. The bag fraction parameter was set to 0.5 to indicate that each tree in the random forest was built using 50% of the total samples randomly selected through bootstrapping (You & Dong, 2020). This adds variability to the training process and further enhances the ensemble's ability to capture diverse patterns. Other parameters, including out-of-bag, were set "true" while variables per split, and seed were kept at their default values at the square root of the number of variables and 0 respectively in the GEE environment (P. Duncan et al., 2023)

2.3.2 Accuracy assessment

The accuracy assessment of the 2023 classified land use/land cover map involved the use of validation samples, which constituted 30% of all the samples, as described in section 2.2.2. To calculate the classification accuracy, a confusion matrix was utilized, which presents a summary of the predicted class labels compared to the actual class labels derived from the validation samples. (Chuvieco, 2020). Additionally, the matrix allows for the computation of various accuracy metrics including overall accuracy (OA), Producer accuracy (PA) and user accuracy (UA) (Chuvieco, 2020). OA is the percentage of correctly classified samples out of the total number of samples in the accuracy assessment. PA refers to the percentage of pixels that are correctly identified as belonging to a specific land cover category. UA represents the percentage of correctly classified pixels for each land cover category. (Chuvieco, 2020).

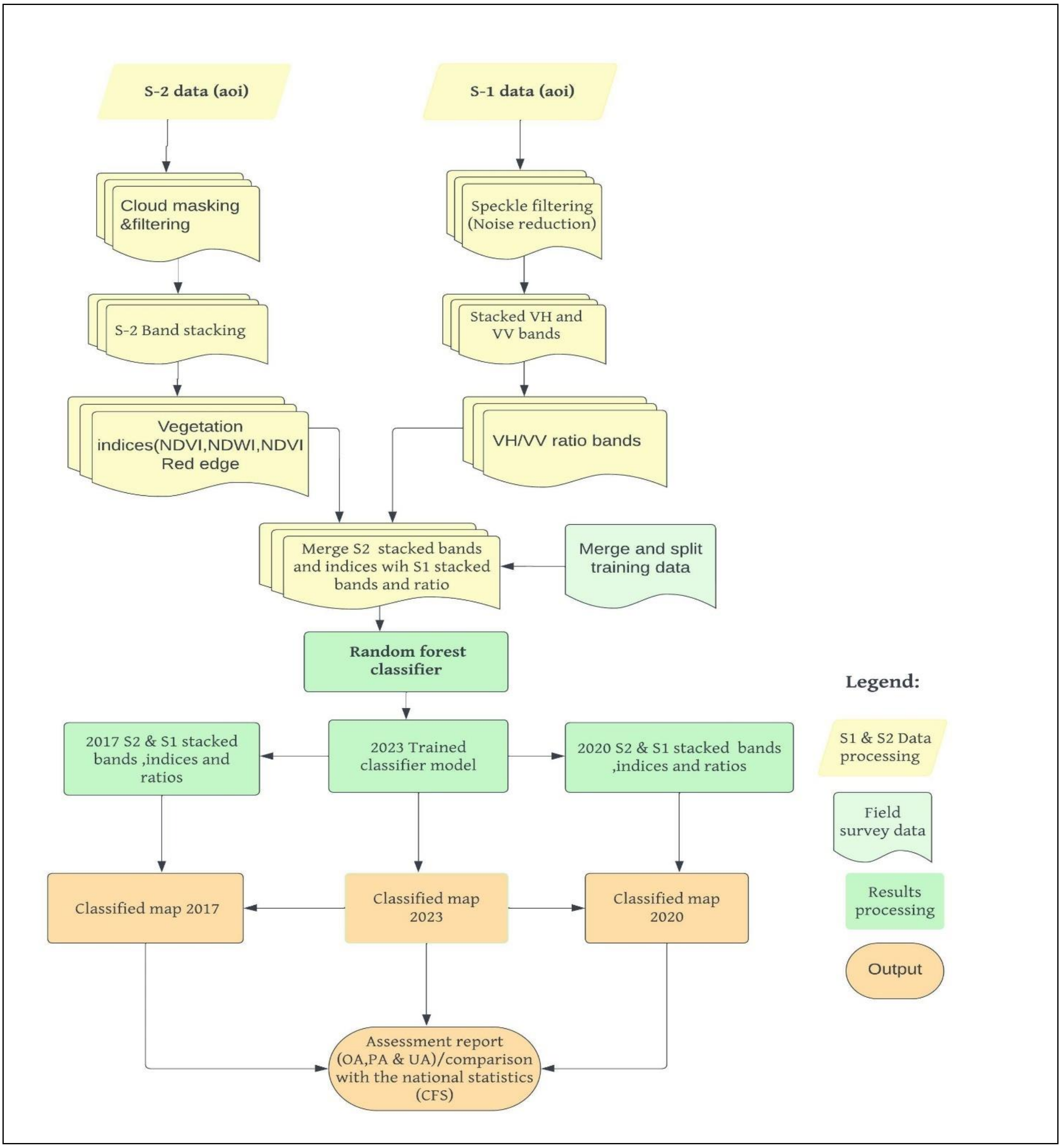


Figure 3. Mapping workflow overview

2.3.3 Classifier transfer to 2017 and 2020

The ground-based reference data was not available for the years 2017 and 2020, which prevented the development of accurate classification models for those years. To overcome this limitation, the RF classifier was trained using the available ground-based reference data for 2023 and applied to the 2017 and 2020 datasets (You & Dong, 2020). This was achieved by replacing the spectral bands and indices of the 2023 classifier with those of the corresponding years while keeping the rest of the classifier unchanged

(Figure *3*). To validate the accuracy of the classification results, statistical results from the crop focus survey (CFS) from the Government of the Republic of Zambia were used (Zam-stats). The CFS provide information on the acreage of each crop type in the study area for the years of interest. The classified maps were compared with the statistics, and the accuracy of the classification was determined by comparing the classified results with the actual acreage of each crop type (Zam-stats).CFS relies on interviews conducted with selected smallholder farmers to gather information on crop cultivation areas. This data collection method involves direct interaction with smallholder farmers, who provide self-reported information on their land use practices.

2.3.4 Food security analysis

To assess food security within households, two scores were utilized: the household dietary diversity score (HDDs) and the household food insecurity access score (HFIAs).(Mango et al., 2014). HDDs is a tool used for assessing food security and nutritional status at the household level, it measures the diversity of foods consumed by household members over a specified period, typically the previous 24 hours (Kuntashula & Mwelwa-Zgambo, 2022). HDDs have been frequently used in surveys, research studies, and monitoring programs to assess dietary diversity at the household level (Mango et al., 2014). It helps identify gaps in dietary diversity, by highlighting vulnerable populations with limited access to diverse food sources and guiding interventions to improve nutrition and food security outcomes (Kuntashula & Mwelwa-Zgambo, 2022). In this study, the HDDs was used to assess the dietary diversity among smallholder farmers. The HDDs is calculated by summing the number of food groups consumed by the household over a reference period, with scores ranging from 0 to 12 (Mango et al., 2014). During the fieldwork, households were requested to report the foods consumed from each food group during the past 24 hours, based on the 12 food groups namely: A, cereals; B, vitamin-rich vegetables; C, roots and tubers; D, dark green leafy vegetables or other vegetables; E, fruits rich in vitamins; F, meat and poultry; G, eggs; H, fish and seafood; I, pulses; J, legumes and nuts; K, milk and milk products; K, oils and fats; and L, sugar and honey (Mango et al., 2014). The count of distinct food groups consumed was tallied and classified into three categories (1-4), medium/moderate dietary (5-8) and high dietary (9-12). The higher the HDDs indicated a greater dietary diversity, reflecting a wider range of nutrients potentially consumed by the household. The calculation of HDDs involves the following procedure:

HDDs (1 to 12) = A + B + C + D + E + F + G + H + I + J + K + L, where A, B, C, D, E, F, G, H, I, J, K, and L represent the 12 food groups (Mango et al., 2018).

The HFIAs is another widely used tool for evaluating food insecurity at the household level. It evaluates the access and availability of food within a household, considering the perspectives and experiences of household members regarding their access to a sufficient and nutritious diet. (Nkembi et al., 2021). The HFIAs provide a quantitative measure of household food insecurity, allowing for comparisons across households or populations (Wambogo et al., 2018). It helps to identify households or communities that are at risk of food insecurity and provides an understanding of the dimensions and severity of food insecurity they experience (Wambogo et al., 2018). The HFIAs consisted of a series of questions related to different aspects of food access and availability. Household heads were asked to respond to each question based on their experiences over the past 30 days (Kolog et al., 2023). In this study, the HFIAs assessed the experiences of smallholder farmers regarding food insecurity across nine key aspects: (Q1) concerns about having enough food; (Q2) consumption of less preferred foods; (Q3) limited variety in food choices; (Q4) inability to consume even less preferred foods; (Q5) inadequate portion sizes of meals;

(Q6) reduced frequency of meals; (Q7) inability to acquire any food; (Q8) experiencing hunger before sleep; and (Q9) going without eating anything for a whole day or night. (Mango et al., 2014). The responses were then scored on a scale of 0 to 27 with higher scores indicating higher levels of food insecurity. The scoring is based on the frequency (F) and severity of food insecurity experiences reported by the household. The scores were subsequently grouped into three categories: food secure, moderately food insecure, and severely food insecure. (Kolog et al., 2023). The HFIAs are calculated by summing the frequency (F) of experiences related to the nine key aspects of food insecurity over the past 30 days.

HFIAs (0 to 27)= (Q1*F*1) +(Q2F2) + (Q3F3) + (Q4F4) + (Q5F5) + (Q6F6) + (Q7F7) + (QF8) + (QF9) (Wambogo et al., 2018).

To analyze the dietary diversity and food insecurity levels among the main stakeholder groups involved in soybean production, namely soybean growers (the primary producers), soybean expanders (ensuring soybean production meets various market demands), soybean sellers (responsible for distribution and marketing of soybean products to consumers), and soybean utilizers (playing a crucial role in addressing food security challenges and promoting a diverse and nutritious diet), the study utilized the box plots as a visual tool for data representation (Nuzzo, 2016). Box plots are graphical representations that display the distribution, central tendency, and variability of a dataset (Nuzzo, 2016). They are useful for comparing multiple groups or variables and identifying potential outliers or skewness in the data, in this case, the major stakeholders in soybean expansions. Box plot analysis allows for the examination of the quartiles, median, range, and potential extreme values of the dataset, providing a valuable understanding of the overall distribution and characteristics of the data (Nuzzo, 2016).

2.3.5 Ordinal logistic regression model

In the last section, the study focuses on the overall factors affecting food security in the district among soybean growers. To analyze the relationship between HDDs, and HFIAs, with the factors affecting food security, ordinal logistic regression was chosen as the appropriate statistical method. Both HDDs and HFIAs are ordinal variables categorized into three levels: low, moderate, and high for HDDs, and food secured, moderately food insecure, and severely food insecure for HFIAs (Lokosang et al., 2011).

Ordinal logistic regression was considered suitable as it allows modelling the probability of an outcome falling into one of the three categories based on the values of independent variables (Appiah-Twumasi & Asale, 2022). Unlike other regression methods, ordinal logistic regression does not assume normality, which is often violated when working with ordinal data (Appiah-Twumasi & Asale, 2022). By using this method, the study assessed the associations between the two dependent variables HDDs and HFIAs and socio-demographic characteristics to identify the most significant predictors of food security in the study area.

The first step involved computing the correlation matrix to assess the relationships between the selected variables (Table 9). Subsequently, highly correlated variables were identified using a correlation cut-off of 0.7. (Soofi, 1990) The resulting highly correlated variables were excluded from the subsequent analyses to mitigate collinearity issues. The regression results (Table 13) highlight the variables that were free from collinearity. Logistic regression models were then fitted for both HDDs and HFIAs using the remaining variables. The coefficients and significance tests were calculated for the fitted models, providing meaningful information about the relationship between the predictors and the outcome variables. These findings were vital in understanding the overall factors influencing food security in the study area.

3 Results

3.1 Land use/land cover classification

Table *5* below presents the overall accuracy, user's and producer's accuracy for the land use and land cover classification in the year 2023. In terms of the UA, the classification achieved an accuracy of 75% for soybean, 80% for maize, 82% for sunflower, 89% for groundnuts, 91% for cotton, 88% for forestry, 95% for build-up, and 93% for water bodies. Regarding the PA, the classification achieved an accuracy of 79% for Soybean, 85% for Maize, 83% for Sunflower, 67% for Groundnuts, 57% for Cotton, 93% for Forestry, 94% for build-up/bare soil, and 85% for Water bodies representing the proportion of pixels correctly identified as belonging to each specific land use/land cover. The land use/land cover classification for 2023 achieved 86%.

Table 5: Overall, user's and producer's accuracy (in %) for the 2023 land use/landcover classification

	2023	
Land use/land cover	PA	UA
Soybean	79	75
Maize	85	80
Sunflower	83	82
Groundnuts	67	89
Cotton	57	91
Forestry	93	88
Build-up/bare soil	94	95
Water bodies	85	93
Overall accuracy	0.86	

Table 6: Comparisons and validation of 2017 and 2020 classification results using CFS (ha)

Year	Soybean	Maize	Sunflower	Groundnuts	Cotton	Forests	Build-up/bare soil	Water bodies
2017	21,890.6	98,875.1	3,113.0	7,451.4	3,182.4	56,821.6	88,760.9	152,760.0
CFS	25,600.8	77,294.6	1,288.1	9,470.2	4,501.8	N/A	N/A	N/A
2020	16,138.8	109,9811.3	5,064.8	11,437.1	7,11.7	40,2519.9	81,128.8	15,7557.9
CFS	14,827.0	83,036.0	3,758.0	9,109.0	5,762.0	N/A	N/A	N/A

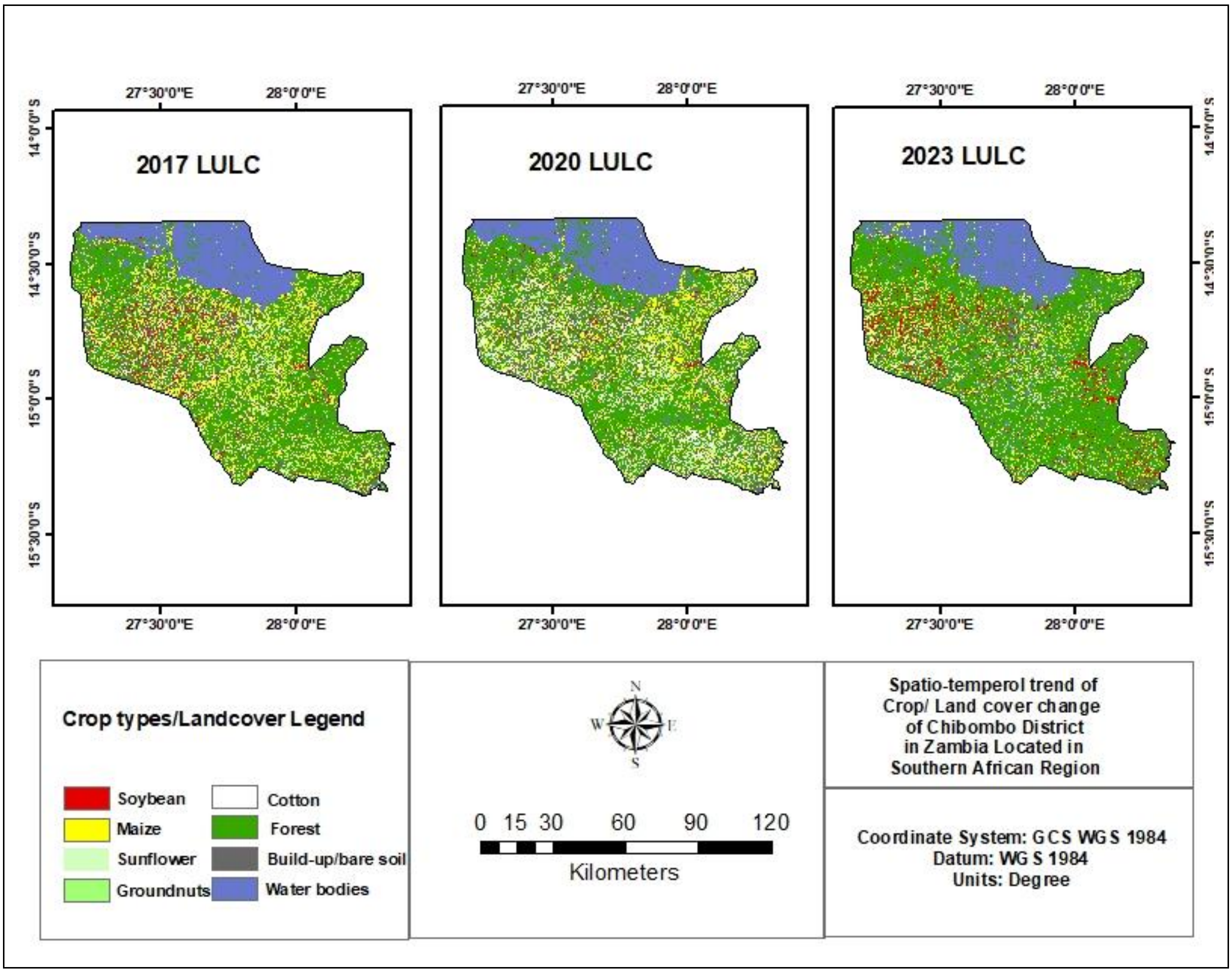


Figure 4 Thematic land cover maps illustrating the soybean expansion between 2017 and 2023

The table below summarizes the areas dedicated to various crop types and land cover categories during the specified years. The categories include soybean, maize, sunflower, groundnuts, cotton, forestry, build-up, and water bodies. The data offers valuable information regarding the changes in land use patterns and the cultivation of specific crops within the study area.

In 2017, soybean cultivation covered an area of 21,890.6 ha, which decreased to 16,138.8 ha in 2020. However, by 2023, the soybean cultivation area increased to 56,169.4 hectares, representing a substantial percentage change of +156.7% from 2017 to 2023 and a remarkable +248.9% from 2020 to 2023. Maize the staple food crop displayed a contrasting trend. The maize area covered 98,875.1 ha in 2017, expanding to 109,811.3 ha in 2020. However, by 2023, the maize area witnessed a significant reduction to 74,640.2 ha, resulting in a percentage change of -24.4% from 2017 to 2023 and a further -32.0% from 2020 to 2023.

Table 7: Crop type/land cover change transitions between 2017 and 2023 (ha)

Year	Soybean	Maize	Sunflower	Groundnuts	Cotton	Forests	Build-up/bare soil	Water bodies
2017	21,890.6	98,875.1	3,113.0	7,451.4	3,182.4	56,821.6	88,760.9	152,760.0

2020	16,138.8	109,9811.3	5,064.8	11,437.1	7,11.7	40,251.9	81,128.8	157,557.9
2023	56,169.4	74,640.2	6,731.3	8,643.7	7,194.6	53,491.4	77,416.7	110,676.0

In terms of other crops, sunflower cultivation experienced a notable increase. The area dedicated to sunflower cultivation was 3,113 ha in 2017, which increased to 5,064.8 ha in 2020, and further expanded to 6,731.33 hectares in 2023. These changes represent percentage increases of +115.9% from 2017 to 2023 and +32.9% from 2020 to 2023, indicating a growing interest in sunflower cultivation among smallholder farmers. Groundnuts, on the other hand, showed a mixed pattern. The area for groundnut cultivation was 7,451.4 hectares in 2017, which increased to 11,437.1 ha in 2020. However, by 2023, the groundnut area decreased to 8,643.7 hectares. This results in a percentage change of +16.0% from 2017 to 2023, but a decrease of -24.4% from 2020 to 2023.

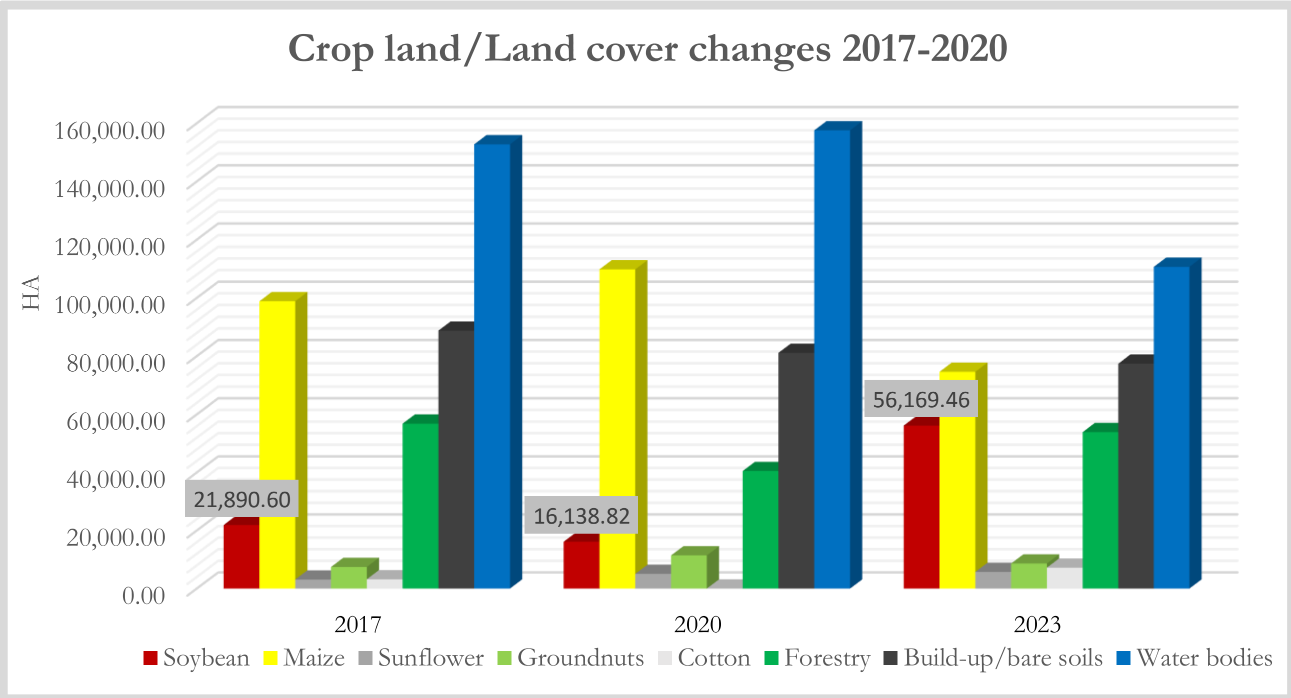


Figure 5. Cropland/landcover changes from 2017 to 2023

In addition to cropland expansions, changes in land cover were also observed. Forest areas displayed a decreasing trend, with 56,821.6 hectares in 2017, declining to 40,519.9 ha in 2020, and then increasing to 53,918.4 ha in 2023. This represents a percentage change of -5.2% from 2017 to 2023 and an increase of +33.2% from 2020 to 2023. Build-up and bare soils showcased a slight decline over the years. The built-up area and bare soils were 88,760.9 ha in 2017, which was reduced to 81,128.8 ha in 2020, and further decreased to 77,416.74 ha in 2023.

3.2 Impact of soybean expansion on other land use/land cover

Between 2017 and 2023, notable changes occurred in the land use/land cover classes. (Table 8). Soybean experienced a substantial increase in area, gaining a total of 34,278.8 ha, which corresponds to a gain of 156.5%. Soybean expanded its area by acquiring land from other classes. The largest contribution to soybean expansion came from maize, with 11,449.7 ha converted to soybean cultivation, accounting for 34.1% of the total gain. Additionally, sunflower contributed 6,816.9 ha (19.9%), groundnuts contributed 288.5 ha (0.8%), cotton contributed 3,095.2 ha (9.0%), forests contributed 11,011.2 ha (32.1%), and build-up/bare soil contributed 13,927.1 ha (40.7%).

Maize cultivation experienced a substantial loss of 24,234.9 ha, representing a decrease of 24.5% in its area. The primary destination for the areas lost from maize was soybean cultivation, which gained 11,449.7 ha from maize, accounting for 47.3% of the total loss. This indicated a significant shift in agricultural practices, with smallholder farmers choosing to convert maize fields to soybean cultivation. Sunflower cultivation witnessed a decrease of 382.3 ha, representing a decline of 12.3% in its area. The majority of the lost areas from sunflower were converted to soybean cultivation, which gained 6,816.9 ha from sunflower, constituting 82.2% of the total loss from sunflower.

Groundnut cultivation experienced a relatively minor decrease of 3.9 ha, corresponding to a decline of 1.3% in its area. The areas lost from groundnuts were distributed across various land cover classes, including soybean, maize, and cotton. However, the extent of the losses from groundnuts was relatively small compared to the gains in other land cover classes. Cotton cultivation also experienced a decrease in the area, losing 641.72 ha, which accounts for a decrease of 20.1%. The areas lost from cotton were primarily gained by soybean cultivation, with 3,095.2 ha transitioning to soybean cultivation, representing 482.2% of the total loss from cotton.

Table 8: Land use/land cover change between 2017 and 2023 (ha).

LULC	Soybean	Maize	Sunflower	Groundnuts	Cotton	Forests	Build-up/bare soil	Water bodies	Total area 2023
Soybean	8,763.40	11,449.70	6,816.90	288.50	3,095.20	11,011.20	13,927.10	817.40	56,169.40
Maize	6,256.28	33,724.40	8,475.80	591.40	3,197.10	24,151.72	11,575.70	2,667.80	74,640.20
Sunflower	670.60	341.10	2,904.20	231.40	809.60	568.60	325.00	880.80	6,731.30
Groundnuts	369.20	2,182.40	748.60	67.50	268.90	4,155.20	717.00	125.90	8,634.70
Cotton	711.36	1,940.20	885.76	54.28	641.72	1,445.39	1,431.39	84.50	7,194.60
Forests	336.46	699.65	731.85	940.42	420.12	33,923.15	7,513.77	8,925.98	53,491.40
Build-u/bare soil	1,287.04	2,088.02	1,714.36	94.84	237.43	8,233.08	757.65	28,047.34	77,416.70
Water	682.01	4,153.87	582.69	20.68	553.08	30,704.07	12,927.95	96,008.59	110,676.00

bodies									
Total (2017)	21,890.6	98,875.1	3,113.0	7,451.4	3,182.4	56,821.6	81,908.94	152,760.0	432,854.40

Forest areas witnessed a loss of 3,330.2 ha, indicating a decrease of 5.9% in their area. The destinations for these lost areas varied, with significant gains observed in the build-up/bare soil class (2,509.1 ha), representing 75.4% of the forest loss, and to a lesser extent in the cotton class (162.4 ha) and other land cover classes. Water bodies experienced a substantial loss of 42,084.0 ha, corresponding to a decrease of 27.6%. The destinations for these lost areas primarily include the build-up/bare soil class (28,963.2 ha), constituting 68.8% of the water bodies loss, and to a lesser extent, the forests class (8,800.1 ha) and other land cover classes.

Overall, soybean, as the dominant land use class, experienced significant gains from various classes, notably maize, sunflower, groundnuts, cotton, forests, and build-up/bare soil. Maize, groundnuts, cotton, forests, and water bodies experienced losses to varying degrees, contributing to the expansion of soybean cultivation. These land cover changes reflect the dynamic nature of agricultural practices in the study area.

3.3 The extent of soybean expansion in Zambia and how has it changed over time?

In 2017, 21,890.6 ha of land was used for soybean production. However, by 2020, the area of land used for soybean cultivation had decreased to 16,138.82 ha. In 2023, the area of land used for soybean cultivation increased significantly to 56,169.46 ha. Overall, results indicate that there was a decline in soybean production between 2017 and 2020, but a significant increase in soybean production by 2023. This translates into a 247.63% increase in soybean area between 2020 and 2023. Comparing the two, it can be concluded that the increase in soybean land use area between 2020 and 2023 is much greater than the decrease in soybean area between 2017 and 2020.

3.4 The impact of soybean expansion on food security among smallholder farmers

The table below provides a summary of the two dependent variables, HDDs and HFIAs, as well as the independent variables used to predict household food security in the study area.

Table 9: Summary description of the variables used in the analysis of food security

Variables	Description and measurement
Dependent variable	
HDDs	Low = 1, Medium=2, High = 3
HFIAs	Food secured= 1, Moderately insecure =2, Severely insecure=3
Independent variable	
Gender(G)	1 =Male household head ; 0 = Female head
Age (A)	Household age (years)
Education	No education =0, Primary= 1, Secondary=2, Tertiary=3
Marital Status (MS)	Single=0, Marriage=1, Divorced =2, Window=3

Household size(HS)	Number of household members
Sources of income (SI)	Farming=1, Labourer=2,Remittances=3
Source of food (SF)	Own+ market=1, Market only=2, Farm+ work =3,Own farm= 4, all sources =5
Distance to the market (DM)	Far=1, Near=0
Distance to the main road (DR)	Far=1, Near=0
Size of household Land(SHL)	Land size in hectares
Assets (AS)	HH own assets=1, HH without assets =0
Employment (EMP)	Formal =1,informal=0
Labour (LR)	Hiring labour=1, 0=Not hiring labour
Livestock (LS)	Own livestock=1, 0= Not owing livestock
Number of crops (NC)	Number of crops grown per HH
Soybean grower(SG)	Growing soybean =1; 0=Not growing
Soybean Expansion (SE)	Expanded soybean =1; 0= not expanded
Soybean Utilization(SU)	Utilizing soybean=1;0=Not utilizing soybean
Input support (IS)	Support=1; 0=Not receiving support
Selling soybean (SS)	Selling soybean =1; 0=Not selling soybean

Table *10* below provides information on the frequencies, percentages, mean, and standard deviation of the different variables used to analyze the factors impacting household food security in the Chibombo district. The mean column represents the average value of each variable within the sampled population, while the standard deviation column indicates the level of variability for each variable among the sampled population.

Table 10 : Sample characterisation

Variable	Frequency	Percentage	Variable	Frequency	Percentage
Gender			Assets		
Male	155	64.6	Tractor/plough/oxen	76	31.7
Female	85	35.4	basic hand tools	164	68.3
Education			Employment		
No education	42	17.5	Formal	16	6.7
Primary	89	37.1	Informal	224	93.3
Secondary	83	34.6	Labour		
Tertiary	26	10.8	Hiring	66	27.5
Marital status			No	174	72.5
Marriage	166	69	Livestock		
Single	18	7.5	No	94	39.2
Divorced	22	9.2	Own	146	60.8
Widow/widower	34	14	Soybean grower		
Income sources			Non-grower	34	14.2
Farming only	124	51.7	Grower	206	85.8
Farming +Labourer	74	30.8	Soybean expansion		
Remittance/Employment	42	17.5	Not expanded	62	25.8
Food sources			Expanded	146	60.8

Own farm +market	214	89.2	Not growing soybean	32	23.3
Market only	1	1.6	Soybean utilization		
Farm+ work for food	16	6.7	Not utilizing	139	57.9
Own fam only	2	0.8	Utilization	55	22.9
Own farm+ market+ external support	4	1.7	Not growing soybean	46	19.2
Market distance			Farm input support		
Near	114	48	No support	142	59
Far	126	52	support	98	41
Distance to road					
Near	154	64			
Far	86	36			

Variable	Mean	Std. Deviation
Land size	6.45	7.999
Age of the household head	49.08	12.507
Number of crops grown	1.84	.709

Table 11 provides an overview of HHDs and HFIAs in the study area. Among the 240 households interviewed, 40% had a low dietary diversity score, 40.4% had a medium/moderate score, and 19.6% had a high score. These findings indicate a significant proportion of households with low or moderate dietary diversity, potentially posing a threat to food security. Additionally, the results show that over half (50.8%) of the households were severely food insecure, while 28.7% were food secure and 20.4% were moderately food insecure. Overall, the combination of HDDs and HFIAs results reveals that a substantial number of households have low or moderate dietary diversity, coupled with a high incidence of food insecurity.

Table 11: Percentage distribution and frequencies of HDDs and HFIAs

HDDs		
Category	Frequency	Percentage
Low dietary diversity score (1-4)	96	40.0
Medium/moderate diversity score (5-8)	97	40.4
High dietary diversity score (9-12)	47	19.6
HFIAs		
Category		
Food secured	69	28.7
Moderate food insecure	49	20.4
Severely insecure	122	50.8
Total	240	100

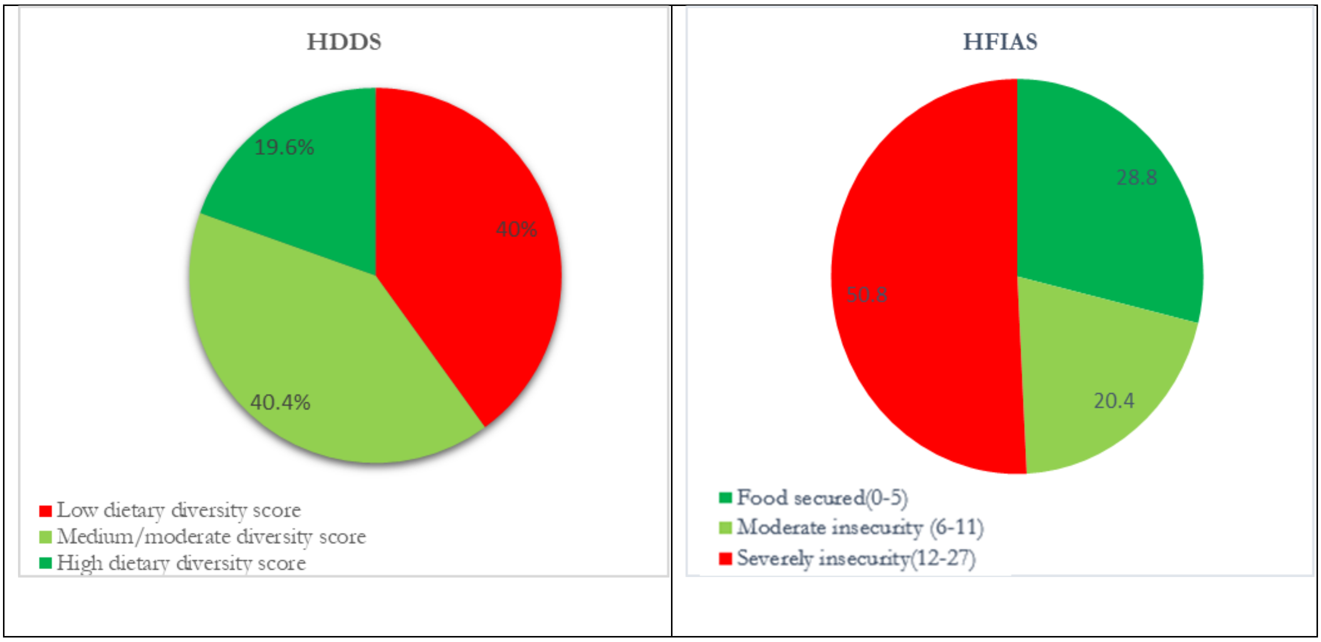


Figure 6. Visual representation of the percentage distribution for both HDDs and HFIAs

3.5 Effects of the expansion on HHDs and HFIAs across the wards

Table 12 presents a summary of the comparison between HDDs and HFIAs across various wards of soybean cultivation areas. It presents data for each ward, including the ward number, the corresponding soybean cultivation area in ha, the percentage of households categorized as having low, moderate, or high HDDs, and the percentage of households categorized as food secure, moderately insecure, or severely insecure based on HFIAs. From the results, variations can be observed across different wards. Some wards have a higher percentage of households with low or moderate levels of hunger, while others have higher percentages of households experiencing severe food insecurity. The wards also differ in terms of their size, with areas ranging from relatively small to larger ones. These findings highlight the diversity in the levels of food security and poverty levels across the wards

Table 12: Comparison of HDDs and HFIAs across the ward concerning the soybean area

WARD No.	AREA (ha)	HDDs (low) %	HDDs (moderate)%	HDDs (high)%	HFIAs (food secure)%	HFIAs (moderately insecure)%	HFIAs (Severely insecure)%
1	1192.94	53.3	33.3	13.3	13.3	13.3	73.3
2	1907.14	53.3	40	26.7	46.7	20	33.3
3	2416.4	33.3	53.3	13.3	20	6.7	73.3
4	4200.48	40	40	20	20	13.3	66.7
5	1021.42	53.3	26.7	20	20	40	40
6	195.78	66.7	20	13.3	13.3	20	66.7
7	2262.52	40	40	20	46.7	13.3	40
8	2238.14	33.3	40	26.7	53.3	13.3	66.7
9	1140.68	53.3	33.3	13.3	26.7	20	53.3
10	135.97	46.7	20	33.3	33.3	13.3	53.3

11	776.4	13.3	60	26.7	20	26.7	53.3
12	5164.28	0	80	20	53.3	33.3	13.3
13	2534.63	66.7	20	13.3	26.7	6.7	66.7
14	5928.28	60	26.7	13.3	13.3	13.3	73.3
15	21340.29	0	80	20	33.3	60	6.7
16	3553.14	33.3	40	26.7	20	13.3	73.3

Figure *7*. Spatial distribution of soybean cultivation area about the HFIAs and HDDs below provides a comprehensive spatial overview of the HHDs and HFIAs across the wards, considering the soybean area to give an insight into the potential impact of the soybean cropland area on food security. The analysis reveals varying percentages of households with diverse dietary patterns and different levels of food accessibility across the wards.

In Ward 1, with a soybean area of 1,192.94 ha, 53.3% of households exhibit low dietary diversity, while 33.3% demonstrate moderate dietary diversity and 13.3% showcase high dietary diversity. Moving on to Ward 2, with a larger soybean area of 1,907.14 ha, a similar pattern emerges. 53.3% of households display low dietary diversity, 40% exhibit moderate dietary diversity, and 26.7% manifest high dietary diversity. Ward 3, covering a soybean area of 2,416.4 ha, illustrates that 33.3% of households have low dietary diversity, 53.3% maintain moderate dietary diversity, and 13.3% with high dietary diversity. Conversely, Ward 4, characterized by a soybean area of 4,200.48 ha, presents a distinct distribution. In this ward, 40% of households experience low dietary diversity, 40% exhibit moderate dietary diversity, and 20% showcase high dietary diversity.

Ward 5, which encompasses a soybean cultivation area of 1,021.42 ha, 53.3% of households possess low dietary diversity, 26.7% demonstrate moderate diversity, and 20% display high diversity, implying a certain degree of dietary variety. Ward 6, despite having a smaller soybean area of 195.78 ha, reveals that 66.7% of households exhibit low dietary diversity, thereby suggesting limited dietary variety. Similarly, Ward 7, which had an area of 2,262.52 ha, exhibits a distribution comparable to that of Ward 4, with 40% of households experiencing low diversity, 40% displaying moderate dietary diversity, and 20% showcasing high dietary diversity. In Ward 8, comprising a soybean area of 2,238.14 ha, observe that 33.3% of households exhibit low dietary diversity, 40% maintain moderate dietary diversity, and 26.7% manifest high diversity.

Ward 9, with a soybean cultivation area of 1,140.68 ha, we find that 53.3% of households showcase low diversity, 33.3% exhibit moderate dietary diversity, and 13.3% display high dietary diversity. In contrast, Ward 10, characterized by the smallest soybean area of 135.97 ha, displays relatively higher dietary diversity, with 46.7% of households exhibiting low dietary diversity, 20% demonstrating moderate dietary diversity, and 33.3% showcasing high dietary diversity. Moving on to Ward 11, which encompasses a soybean area of 776.4 hectares 13.3% of households maintain low dietary diversity, 60% display moderate dietary diversity, and 26.7% exhibit high diversity. Ward 12 stands out with one of the largest soybean cultivation areas of 5,164.28 ha. Interestingly, no households in this ward possess low diversity. Instead, 80% of households fall under the moderate dietary diversity category, while only 20% belong to the high dietary diversity category.

Ward 13, which encompasses a soybean area of 2,534.63 ha had 66.7% of households exhibit low diversity, indicating a lower level of dietary variety. Similarly, Ward 14, characterized by a soybean area of

5,928.28 hectares, demonstrates that 60% of households possess low dietary diversity, thereby indicating limited dietary variety. Ward 15, which had the largest soybean area of 21,340.29 ha, surprisingly presents no households with low dietary diversity. Instead, 80% fall under the moderate dietary diversity category. Lastly, Ward 16, with a soybean area of 3,553.14 hectares, displays a distribution of 33.3% low diversity, 40% moderate diversity, and 26.7% high diversity among households

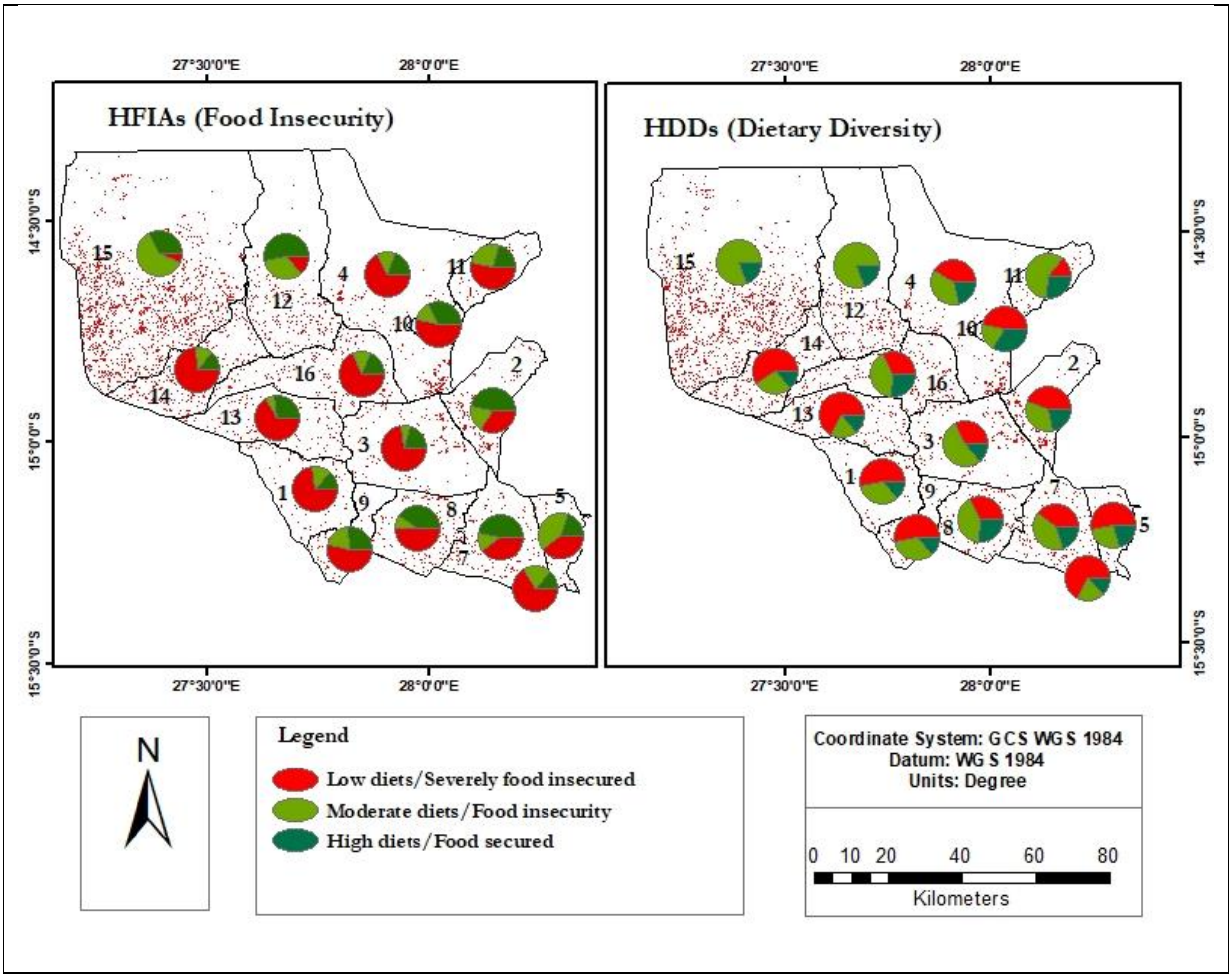


Figure 7. Spatial distribution of soybean cultivation area about the HFIAs and HDDs

Regarding HFIAs, the distribution of moderately insecure and severely insecure HFIAs varies among the different wards. In Ward 1, which encompasses a soybean cultivation area of 1,192.94 ha, 13.3% of households experience severe food insecurity. Moving on to Ward 2, with a larger soybean area of 1,907.14 ha, 33.3% of households face moderate food insecurity. In Ward 3, covering 2,416.4 ha of soybean cultivation, a significant 73.3% of households are classified as severely food insecure. In comparison, Ward 4, characterized by a soybean area of 4,200.48 ha, exhibits severe food insecurity in 66.7% of households.

Ward 5, which encompasses an area of 1,021.42 ha. In this ward, 40% of households are classified as moderately food insecure, while another 40% face severe food insecurity. Surprisingly, despite its smaller soybean area of 195.78 ha, Ward 6 has a high percentage of households (66.7%) experiencing severe food insecurity. Examining Ward 7, with a soybean area of 2,262.52 ha, we find that 40% of households are moderately food insecure. In Ward 8, covering 2,238.14 ha of soybean cultivation, a significant 66.7% of households face severe food insecurity.

Ward 9, with a soybean cultivation area of 1,140.68 ha, 53.3% of households fall under the category of moderately food insecure. In Ward 10, with the smallest soybean area of 135.97 ha, a striking 53.3% of households experience severe food insecurity. Ward 11, covering a soybean area of 776.4 hectares, has 53.3% of households classified as severely food insecure. Remarkably, despite its substantial soybean cultivation area of 5,164.28 ha, Ward 12 exhibits a relatively low percentage (13.3%) of households experiencing moderate food insecurity.

Ward 13, encompassing a soybean area of 2,534.63 ha, has 66.7% of households facing severe food insecurity. In Ward 14, with a soybean area of 5,928.28 hectares, a high percentage (73.3%) of households are classified as severely food insecure. Finally, in Ward 15, boasting the largest soybean area of 21,340.29 ha, only a small portion (6.7%) of households are considered severely food insecure. Ward 16, covering a soybean area of 3,553.14 hectares, displays a high percentage (73.3%) of households facing severe food insecurity.

Both results show that there is no clear trend indicating that wards with larger soybean areas consistently had better HDDs and HFIAs. The HDDs levels varied across different wards regardless of soybean area, with some wards showing higher proportions of low, moderate, or high dietary diversity. For example, Ward 12 had a large soybean area but relatively lower levels of HDDs and HFIAs, suggesting a potential mismatch between soybean production and household food security indicators in this ward, while Ward 10 had a relatively small soybean area of 135.97 ha. However, it had a high HDDs with 33.3% of households falling into the high dietary diversity category. In terms of HFIAs, Ward 10 had a moderately insecure level with 53.3% falling into the moderately insecure category. This further suggests that soybean production may not be the determining factor for household food security outcomes in this ward.

3.6 Effects of the expansion on the dietary diversity among smallholder farmers

Figure *8*. HDDs for the main stakeholders in the soybean expansions above present the analysis of HDDs across different categories, including soybean growers, expansion status, selling soybean, and utilization status.

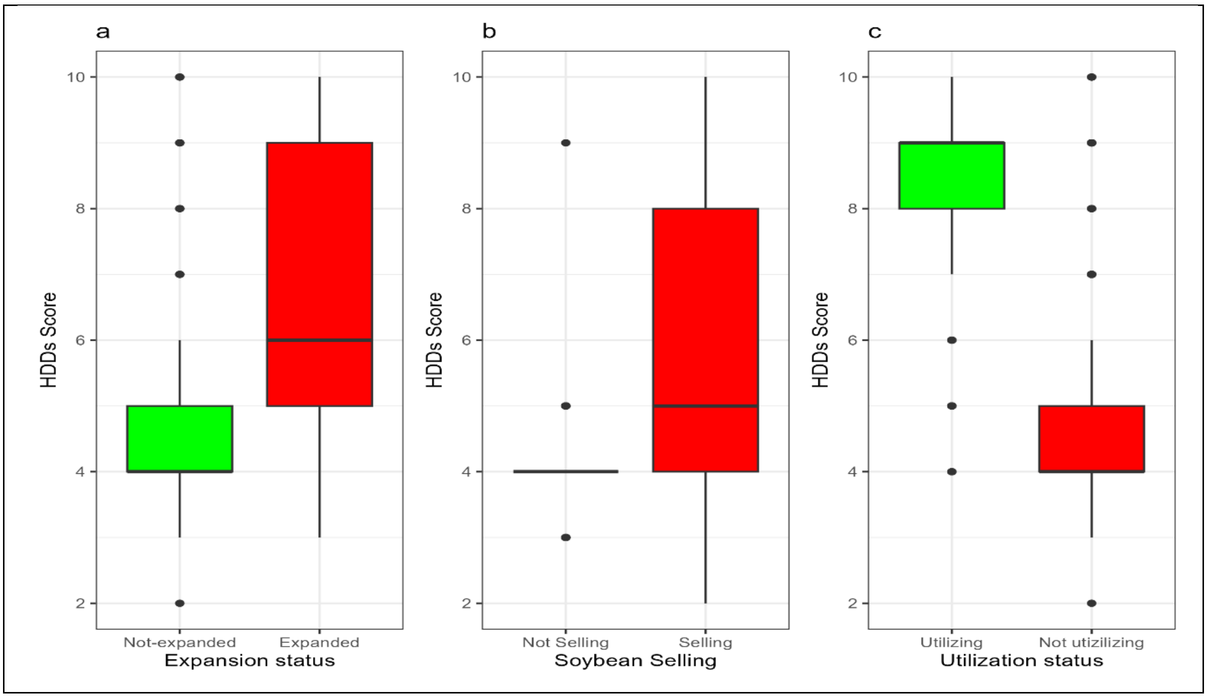


Figure 8. HDDs for the main stakeholders in the soybean expansions

The box plot (a) compares the HDDs between the households that expanded and did not expand the soybean. The households that did not expand the soybean area show a lower quartile score of 4, indicating that 25% of households have relatively lower dietary diversity levels. In contrast, the households who expanded the soybean area demonstrated a higher median score of 6, suggesting improved dietary diversity. Notably, households that expanded the soybean area displayed more consistent HDDs, as there are no outliers. The upper quartile score of 5 for the households that did not expand the soybean area and a 9 for the households that expanded the soybean reflects different ranges of scores. The presence of outliers with high scores for households that did not expand suggests that there are households within this category that exhibit high levels of dietary diversity, even without being part of the soybean expansion.

Box plot (b), indicates that the households who were involved in selling soybean had a slightly higher median score of 5, suggesting relatively better dietary. In contrast, the households who were not involved in the selling of soybean had a median score of 4, indicating lower dietary diversity levels. The lower quartile score of 4 for the households not involved in the selling of the suggests a significant proportion of households in this group have lower dietary diversity while the presence of the outliers indicated that some households exhibited unexpectedly high or low dietary diversity levels.

Box plot (c) reveals that households that were utilizing the soybeans after the harvest had a significantly higher median score of 9, suggesting a higher level of dietary diversity among the surveyed households.

The households that were not utilizing the soybean had a lower median score of 4, indicating a lower level of dietary diversity. The lower quartile score of 4 suggests that 25% of households in this group have relatively lower dietary diversity scores. Additionally households not utilizing the commodity indicate significant variability in dietary diversity with some exhibiting both lower and higher scores.

3.7 Impact of the expansions on the food insecurity among smallholder farmers

Figure 9, below illustrates the different HFIA scores for the major stakeholders. The box plot (a) shows that households that did not expand their soybean cultivation had a higher median quartile score of 13, indicating a relatively elevated level of food insecurity. This is further supported by the upper quartile score of 15, which confirms the prevalence of food insecurity among these households. The presence of outliers suggests that certain households within this group experience exceptionally high levels of food insecurity. In contrast, households that expanded their soybean cultivation demonstrate a median score of 6, suggesting a comparatively lower level of food insecurity. The lower quartile score of 0 indicates that a substantial portion of households that expanded soybean cultivation experienced minimal or no food insecurity, suggesting a consistent pattern of food security. Overall, the results of plot (a) reveals that households who expanded their soybean cultivation were associated with variations in household food insecurity, while those who did not expand exhibited higher levels of insecurity.

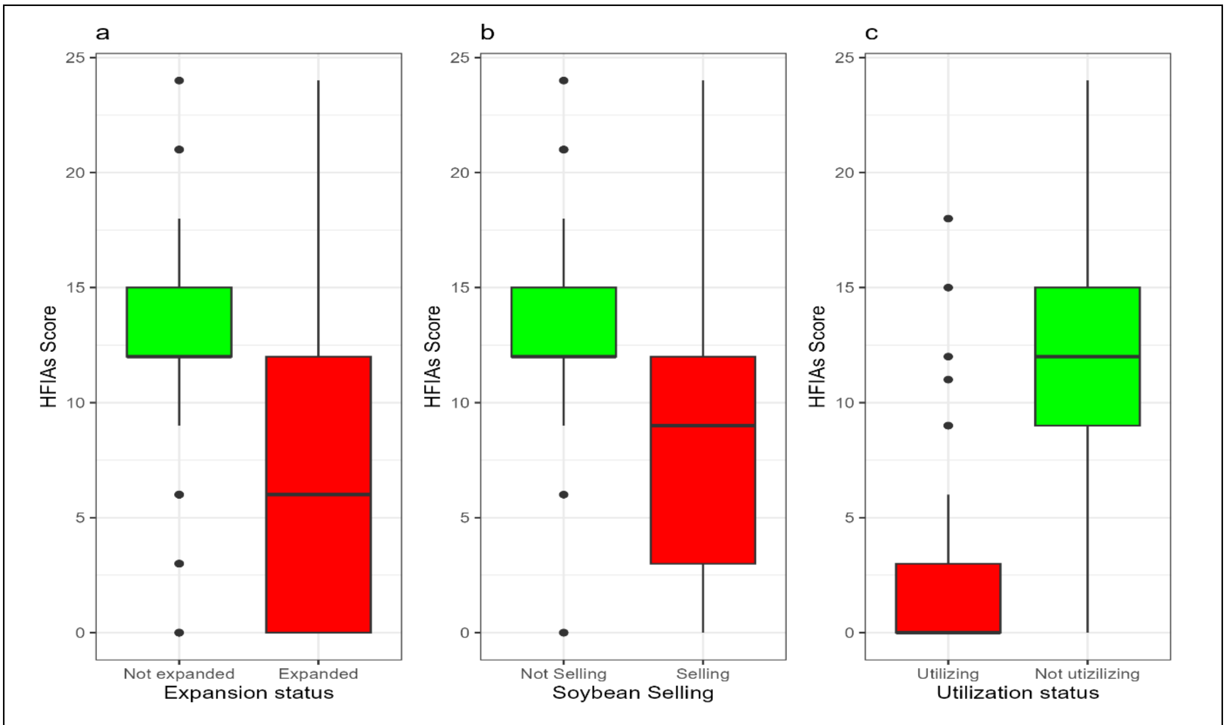


Figure 9 HFIAs for the main stakeholders in the soybean expansion

Box plot (b), focuses on the households who were involved in either selling or not based on HFIAs. The households that did not sell the soybean had a median quartile score of 11, indicating a higher level of food insecurity. Within the same category, some households experienced low levels of insecurity while others experience severe levels of food insecurity. On the other hand, the households that sold the soybean demonstrate a lower median score of 7, implying a comparatively lower level of food insecurity.

The lower quartile score of 4 indicates that a significant proportion of households engaged in selling soybean experienced relatively lower levels of food insecurity.

Box plot (c) compares the households that utilized the soybean and those that did not after. The households that utilized soybean exhibits a median quartile score of 0, suggesting a comparatively lower level of food insecurity. The households that were not utilizing the soybean shows a higher median score of 12, suggesting a comparatively higher level of food insecurity. The absence of outliers indicates a more consistent pattern of food insecurity. Therefore, results for box plot c suggest that the utilization of the soybean may be associated with a reduction in food insecurity, as indicated by the lower median score.

Across all categories, common elements such as the interquartile range, median values, presence of outliers, and potential disparities in both HDDs and HFIAs scores provided valuable insights into the distribution, central tendency, and variations in dietary diversity and food insecurity within each category. The overall analysis also revealed the presence of outliers within each category, indicating further variability in the scores. This suggests that households within different categories exhibited different levels of food insecurity, emphasizing the diversity and heterogeneity in food security levels among these categories.

3.8 Results of the ordinal regression model analysis

The results from the ordinal logistic regression (Table 13) reveal the statistically significant factors affecting food security among major stakeholders in soybean expansions (growers, expanders, and utilization). Education, hiring labour, the number of crops grown per household, and access to inputs are identified as significant factors for both HDDs and HFIAs. Higher levels of education among soybean growers were associated with improved dietary diversity (coefficient = 0.4053011, p = 0.003812), while better access to inputs positively influences dietary diversity (coefficient = 0.8671767, p = 3.723e-05). Higher education levels are also linked to lower food insecurity (coefficient = -0.5343815, p = 0.0002014). The practice of hiring labour is significantly related to reduced food insecurity access (coefficient = -1.0911419, p = 4.016e-06), and an increased number of crops cultivated and improved access to inputs have positive effects on food insecurity access (coefficients = 0.4859041, p = 0.0075451, and -0.4707135, p = 0.0239902 respectively).

Table 13: Ordinal logistic regression results

Variable	HHDDS		HHFIAS	
	Coefficient	Pr(>|z|)	Coefficient	Pr(>|z|)
Gender	-0.2878645	0.239566	-0.0097308	0.9674622
Age	0.0047006	0.550202	-0.0064085	0.4121838
Education	0.4053011	0.003812 **	-0.5343815	0.0002014 ***
Marital status	-0.2363732	0.124641	0.2232187	0.1264030
Household size	0.0349603	0.220686	-0.0615267	0.0480498
Income sources	0.1466931	0.280571	-0.2408934	0.0849670
Market distance	0.1771788	0.331867	-0.3228190	0.0827706
Hiring labour	1.2391075	1.113e-06 ***	-1.0911419	4.016e-06 ***
Livestock	0.3424912	0.131283	-0.1910155	0.4056151
Number of crops	-0.9639665	6.792e-07 ***	0.4859041	0.0075451 **

Soybean grower	0.6848559	0.240163	0.3689400	0.5588338
Expanded soybean	0.1158732	0.511371	-0.2903240	0.1029830
Soybean utilization	-0.1611603	0.380816	0.2037868	0.2738129
Access to inputs	0.8671767	3.723e-05 ***	-0.4707135	0.0239902 *
Soybean selling	-0.1418157	0.775071	-0.8146898	0.1619621

Significant codes: 0 '***' 0.001 '**' 0.01 '*'

4 Discussion

4.1 Soybean cropland expansion

In this study, remote sensing data from S1 and S2 satellites were utilized in the mapping of the expansion of soybean, along with ground truth data, to train and validate the classification results for the land use/ land cover. The RF algorithm was used for the classification, resulting in highly accurate maps depicting the expansion of soybean cropland from 2017 to 2023 achieving an overall accuracy of 86% for the year 2023, revealing a significant increase in soybean production between 2020 and 2023, with a remarkable 247.63% increase in cultivated area. The 2017 and 2020 classifications were validated using the crop focus survey (CFS) with no significant differences. The study revealed that soybean is the dominant crop in terms of its influence on other food crops and land cover changes, specifically in terms of gaining a larger area from other classes in support of the studies of Savala et al., (2022) Indicating that soybean expansion is primarily driven by the conversion of land previously used for other purposes to soybean cultivation. Their research highlights how land is being transformed from its original land use to accommodate the cultivation of soybeans. This study provides evidence supporting their studies due to the substantial increase in soybean cultivation during the same period. The total area dedicated to soybeans expanded by 34,278.8 ha, representing a significant gain of 156.5%. This expansion was due to the conversion of land from other land use/land cover classes. The largest contribution to soybean's expansion came from maize a country's staple food crop, with 11,449.7 ha converted to soybean cultivation, accounting for 34.1% of the total gain. Additionally, sunflower contributed 6,816.9 ha (19.9%), groundnuts contributed 288.5 ha (0.8%) and cotton contributed 3,095.2 ha (9.0%). This significant land conversion from other food crops is attributed to the high demand for soybean, a protein-rich commodity mainly for stock feeds as highlighted in the studies by Voora et al., (2020). The findings of another study conducted by Phiri et al., (2019) align with this study, as both results indicate a high demand for forests, which are considered to be more fertile, leading to increased soybean expansions among smallholder farmers. In this study, forests accounted for 11,011.2 ha (32.1%) of the total land conversion, while build-up/bare soil contributed 13,927.1 ha (40.7%).

Although both soybean cultivation area and production have shown fluctuations since 2005 (Figure *1*) the overall trend has been towards expansion, with more instances of increased cultivation compared to decreases on average. However, these fluctuations are not primarily driven by changes in demand but rather by climate variability, as highlighted in a study conducted by Chilambwe et al. (2022). Their research focused on modelling the relationship between climate variables and crop yields, revealing that climate variables played a crucial role in influencing both maize and soybean yields. They assessed the potential impacts of future climate change, the study simulated mean yield changes for maize and soybean under different climate scenarios over a 20-year period. The projections indicated that soybean cultivated area and production were fluctuating while other crops such as maize cropland exhibited both increases and marginal variations in many districts of the central province including Chibombo. Their findings provide valuable information for understanding other factors contributing to the dynamics of soybean and their vulnerability to climate variability and change. Similarly, the findings of this study revealed similar fluctuations in the soybean cropland area. Between 2017 and 2020, there was a decrease in soybean cultivated area by 26.3% However, from 2020 to 2023, there was a significant increase in soybean cultivation, with a percentage increase of 247.4% in soybean cultivation.

4.2 The impact of the soybean expansion on dietary diversity and food insecurity

Similar to other African countries, the expansion of soybean areas by smallholder farmers in Zambia has been actively encouraged to promote crop diversification as a way to enhance farmers' cash income and improve nutritional security (Kapulu et al., 2023). However, the findings of this study suggest that the expansion of soybean cultivation has a limited impact on increasing dietary diversity and improving food security among smallholder farmers engaged in soybean cultivation. Farmers are primarily driven to expand the soybean for commercial purposes, focusing on its economic potential rather than considering it as a staple food crop for household consumption as revealed by the studies published by Kapulu et al., (2023). Their studies suggested that this could be attributed to the emphasis placed on promoting soybean as a cash crop than a food crop in the messaging and campaigns. The promotional efforts and messaging surrounding soybean cultivation have primarily highlighted its potential economic benefits, which may have influenced smallholder farmers to prioritize expanding cropland for commercial purposes contrary to considering the commodity as a significant food crop for household consumption. Another study by Siamabele, (2019) stated that there has been an increase in soybean cropland area over the past two decades, primarily for the production of livestock feeds and edible oils. This exponential growth in soybean production is fuelled by the increasing livestock sectors in the SAR, which has facilitated increased exports of soybean products such as stock feeds.

The continuous demand for soybean products has motivated smallholder farmers to invest in soybean production through expansions of their soybean croplands resulting in more expansions. Although a recent study by FAO food balance sheets reveals that soybean has made significant contributions to the availability of essential dietary nutrients such as calcium, protein, energy, and iron at the national level in Zambia (Kapulu et al., 2023). However, this study shows that soybean expansion did not contribute to more diverse diets at the household level, instead, the findings revealed a significant proportion of households engaged in soybean expansions with low or moderate dietary diversity, posing a threat to food security. Additionally, the results show that over half of the households surveyed were severely food insecure while 20.4% were moderately food insecure and only 28.7% were food secure. This is likely because soybean is processed into secondary products such as oils and livestock feeds, rather than directly consumed. The low consumption of meat products and poultry suggests that there is little contribution coming from soybean as animal feed to the diets of households according to the study of Chianu et al., (2009). conversely, meat and poultry products utilizing soybean products are often more costly and are typically processed in urban areas, often located far away from rural regions where smallholder farmers reside (Kapulu et al., 2023). This geographical barrier limits the accessibility of soybean by-products for many households, thereby posing threats to food security. As a result, the limited availability and higher costs associated with soybean by-products may hinder the ability of rural communities to access these nutritious food options as stated in the studies of Chianu et al., (2009). This statement aligns with the findings of this study that have reported lower scores for both HDD and HFIA among households that did not utilize soybean as a food source. Kapulu et al., (2023), in their studies, recommended the need to train smallholder farmers in the domestic processing of soybean to encourage its consumption and utilization and further proposed that these challenges should be done in conjunction with existing intervention programs, such as the Scaling Up Nutrition Technical Assistance (SUN TA) program, which aims to promote the consumption of nutritious food by utilizing locally available food in Zambia.

4.3 Factors affecting food security among soybean farmers

The results of the ordinal logistic regression analysis highlight significant factors influencing food security among key stakeholders involved in soybean expansions, including growers, expanders, and utilization. Among these factors, education, labour hiring practices, the number of crops cultivated per household, and access to inputs were identified as statistically significant variables affecting both food security.

This study examined different levels of education among household heads, including, no education, primary, secondary and tertiary education. These various education categories were tested to explore their impact on the food insecurity status of households. The results suggest that higher levels of education were associated with higher household dietary diversity scores and lower levels of household food insecurity access scores. Households headed by individuals with limited educational attainment were found to have poor diets and were more vulnerable to food insecurity, in contrast to households headed by individuals with higher levels of education. The findings of this study are in line with the research conducted by Assefa & Abide, (2023), which similarly highlighted the significance of education in enhancing knowledge, awareness, and decision-making skills concerning recommended agronomic practices and food choices. Hiring labour exhibited highly significant coefficients for both HDDs and HFIAs, indicating that households that hire labour have a higher likelihood of achieving higher dietary diversity and lower levels of food insecurity access scores. Hiring labour provides smallholder farmers with additional labour enabling them to expand their agricultural land and produce a variety of crops. Schmitt-Harsh et al., (2020), conducted a study that also yielded similar results, emphasizing the positive influence of labour hiring on food security outcomes.

The number of crops grown by the household had a higher significance on the diets, the higher number of crops grown by the household, the better the diets. Growing a diverse range of crops increases the availability of different types of food crops, contributing to a varied and balanced diet. The findings of Mango et al., (2014) support this study's results, demonstrating a positive relationship between crop diversity and dietary diversity. Their research suggested that smallholder farmers who engaged in crop diversification were able to enhance their income through the sale of cash crops that will not only boosts their financial situation but also enhances food security by enabling them to consume their products and purchase additional food crops with the income earned from cash crop sales. Access to inputs showed a significant coefficient meaning that Improved access to agricultural inputs, such as certified seeds, fertilizers, and weed management, was associated with higher dietary diversity scores and lower levels of food insecurity. According to Tuni et al., (2022), access to inputs enables farmers to enhance their agricultural production, leading to food availability and improving household food security situation.

Furthermore, this study uncovered variations in dietary diversity and food insecurity outcomes among different wards, regardless of the extent of soybean cultivation in those areas. The absence of a clear relationship between soybean cultivation areas and dietary diversity and food security outcomes suggests that factors beyond soybean production play a significant role in household food security. A related study conducted by Li et al., (2021) highlighted that the size of land owned by smallholder farmers is not the sole determinant of improved food security. Instead, productivity emerged as the main driver of food security. Their study further revealed that many smallholder farmers have been expanding their cropland but with low yields. The study by Cornelius & Goldsmith, (2019) indicated that although the soybean area has expanded in Africa, the levels of production are low among smallholder farmers. This limited production has resulted in minimal contributions to improved food security in the region. This finding is consistent with the results of this study, particularly regarding the fluctuating soybean yields alongside the expansion of cultivated areas (Figure *1*). The central province of Zambia where Chibombo district is located is one of the provinces that has been experiencing decreasing yields from 2.4 tonnes per hectare to (t/ha) in 2013 to 1.3 t/ha in 2023 (Zam-stats). The study conducted by Mango et al., (2014) identified several factors contributing to the low yields experienced by smallholder farmers, including high poverty levels, limited income, and inadequate access to inputs and equipment directly impacting their ability to achieve higher yields even when they attempt to expand on soybean cultivated areas.

Other important findings of this research are the variations in terms of HDDs and HFIAs among the different stakeholder categories involved in soybean production. These variations observed are similar to the variations in the findings of Kapulu et al., (2023) indicating that most smallholder farmers often regard soybean as a means to generate income rather than a staple food for consumption. Consequently, many farmers choose to sell all their soybean produce, resulting in limited access to other food products. Their studies further pointed out that geographical factors, such as long distances to markets, can further hinder their ability to obtain alternative food sources. Additionally, Li et al., (2021), revealed that even if some smallholder farmers expand their cash crops such as soybean, the overall productivity remains low thereby resulting in variations in dietary diversity and food insecurity among different stakeholders.

4.4 Implications and limitations of the study

The classification results obtained in this study were subject to certain uncertainties. The process of classifying land use/land cover using remote sensing data was influenced by various factors, including the quality of the data itself, limitations in image resolution due to cloud cover and shadows, the selection of classification algorithms, and the availability of ground truth data for validation. The collection of the ground truth data was challenging due to limited accessibility and flooding, which may have impacted the accuracy of the classification. These uncertainties could have led to errors or misclassifications in the final land use/land cover maps. Consequently, it is important to highlight these limitations and view the classification results as approximations rather than exact representations of the true land cover and land use conditions.

The study was conducted at a specific time point during the farming season, and it is important to note that dietary patterns and nutrient availability may vary during leaner months. Seasonality plays a significant role in influencing dietary patterns and food availability, which in turn affects the availability of dietary nutrients and overall nutrient adequacy. Thus, future studies could design to capture the factors affecting food security at the beginning of the farming season and after harvesting to have a comprehensive understanding of the dynamics that affect food security throughout the year. For example, food supply among smallholder farmers in rural settings can be significantly influenced by factors such as seasonality and geographical characteristics, including the proximity to food markets. Seasonal variations in agricultural production can impact the availability and diversity of food items, affecting the overall food supply in rural areas. Additionally, the geographical features of an area, such as its distance from food markets, can impact the accessibility and affordability of food for rural communities. These factors need to be taken into consideration when assessing and addressing food security issues.

The other limitation is the low number of household samples captured during the survey. This limitation stems from practical challenges and constraints in data collection, such as limited resources, time, and accessibility to some households. As a result, the small sample size may not fully represent the diverse range of households and their specific food security situations within the study area. The limited sample size may restrict the generalization of the findings and may not capture the full heterogeneity and complexity of the relationship between soybean expansion and household food security in the district. Therefore, the study recognizes the importance of increasing the sample size in future research to obtain a more representative sample to have a comprehensive understanding of the dynamics between soybean expansion and food security outcomes.

The findings of this study carry significant implications for smallholder farmers who depend heavily on agriculture as their primary source of income. Relying solely on income generated from soybean cultivation through expansions may not be sufficient for smallholder farmers to overcome poverty and meet their food needs (Kapulu et al., 2023). Additionally, the low productivity of small-scale soybean farming hinders the ability to earn a sustainable income from soybean cultivation (Nkonde et al., 2021). To address the food security challenges associated with soybean cropland expansions, it is crucial to implement additional interventions that complement the income pathway. One such intervention is providing nutrition education and programs to farmers, enabling them to make informed decisions about how to utilize their agricultural income, including income from soybean production, in purchasing nutrient-rich foods (Kapulu et al., 2023). Furthermore, subsidizing farm implements such as tractors and ploughs can enhance the productivity of smallholder farmers, making their agricultural activities more efficient and profitable (Kapulu et al., 2023). Promoting crop diversification is another important strategy that can contribute to improving overall food security among smallholder farmers, as it reduces their dependence on a single crop and increases the variety and availability of food resources. By implementing these interventions, it is possible to address food security challenges faced by smallholder farmers who have invested in soybean production through the expansions, thereby enhancing their overall well-being and contributing to sustainable agricultural development.

5 Conclusion and recommendation

The study successfully used satellite imagery, ground reference data, and the RF algorithm to accurately map soybean cropland expansion in the Chibombo district of Zambia located in the southern region. The results from the classification results showed that soybean has significantly expanded from 2017 to 2023. Although smallholder farmers have rapidly expanded their soybean cropland to improve food security status, the study reveals that this expansion does not seem to directly contribute to more diverse diets or improved food security at the household level. The findings of the study indicate that a considerable number of households in the Chibombo district have low to moderate dietary diversity, as reflected in the results for HDDs. Additionally, the study reveals a high prevalence of food insecurity among smallholder farmers in the study area, as indicated by the HFIAs results. The observed variations among the major stakeholders involved in soybean cultivation provide additional evidence that soybean expansions alone are not the determining factor for food security among smallholder farmers.

Policymakers and advocates of soybean expansions should take into account socio-demographic factors such as education, crop diversification, and farm input support programs as crucial drivers for enhancing food security at the household level. Policies aimed at improving food security should be supplemented with additional interventions that focus on enhancing soybean utilization within households. This study provides a valuable foundation for informing nutrition-sensitive agriculture policies, specifically about the impact of soybean agricultural expansions on the food security of smallholder farmers. By understanding the implications and challenges associated with these expansions, policymakers can develop effective strategies to enhance food security among smallholder farmers.

6 References

Appiah-Twumasi, M., & Asale, M. A. (2022). Crop diversification and farm household food and nutrition security in Northern Ghana. *Environment, Development and Sustainability*, *0123456789*. https://doi.org/10.1007/s10668-022-02703-x

Assefa, T., & Abide, E. B. (2023). Determinants of food insecurity in rural households: A case of lemo district, southern Ethiopia. *Heliyon*, *9*(1), e12764. https://doi.org/10.1016/j.heliyon.2022.e12764

Awoyemi, A. E., Issahaku, G., & Awuni, J. A. (2022). Drivers of Household Food Security: Evidence from the Ghana Living Standards Survey. *SSRN Electronic Journal*, *13*(October 2022), 100636. https://doi.org/10.2139/ssrn.4252132

Awoyemi, A. E., Issahaku, G., & Awuni, J. A. (2023). Engagement in farming l P Determinants of Household Food Security of. *Journal of Agriculture and Food Research*, 100636. https://doi.org/10.1016/j.jafr.2023.100636

Bégué, A., Leroux, L., Soumaré, M., Faure, J. F., Diouf, A. A., Augusseau, X., Touré, L., & Tonneau, J. P. (2020). Remote Sensing Products and Services in Support of Agricultural Public Policies in Africa: Overview and Challenges. *Frontiers in Sustainable Food Systems*, *4*(May). https://doi.org/10.3389/fsufs.2020.00058

Belgiu, M., & Csillik, O. (2018). Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis. *Remote Sensing of Environment*, *204*(September 2017), 509–523. https://doi.org/10.1016/j.rse.2017.10.005

Blickensdörfer, L., Schwieder, M., Pflugmacher, D., Nendel, C., Erasmi, S., & Hostert, P. (2022). Mapping of crop types and crop sequences with combined time series of Sentinel-1, Sentinel-2 and Landsat 8 data for Germany. *Remote Sensing of Environment*, *269*(April 2021). https://doi.org/10.1016/j.rse.2021.112831

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Bwalya, B. (2022). An evaluation of the agro-dealers ' role in enhancing diversification among farmers of Chibombo district of Zambia An evaluation of the agro- dealers ' role in enhancing diversification among farmers of Chibombo district of Zambia . April.

Castell, G. S., Rodrigo, C. P., de la Cruz, J. N., & Bartrina, J. A. (2015). Escalas de evaluación de la inseguridad alimentaria en el hogar. *Nutricion Hospitalaria*, *31*, 272–278. https://doi.org/10.3305/nh.2015.31.sup3.8775

Chianu, J. N., Ohiokpehai, O., Vanlauwe, B., Adesina, a, De Groote, H., & Sanginga, N. (2009). Promoting a Versatile but yet Minor Crop : Soybean in the Farming Systems of Kenya. *Journal of Sustainable Development in Africa*, *10*(4), 324–344.

Chilambwe, A., Crespo, O., & Chungu, D. (2022). Climate change impacts on maize and soybean yields in Zambia. *Agronomy Journal*, *114*(4), 2430–2444. https://doi.org/10.1002/agj2.21092

Chuvieco, E. (2020). Accuracy Assessment. *Fundamentals of Satellite Remote Sensing*, 356–383. https://doi.org/10.1201/b18954-10

Cornelius, M., & Goldsmith, P. (2019). Soybean yield in Africa. *African Journal of Food, Agriculture, Nutrition and Development*, *19*(5), 15169–15172. https://doi.org/10.18697/ajfand.88.SILFarmDoc12

Doumit, J., & Kiselevm, E. (2017). Comparison of Divert Vegetation Indices for Agricultural Mapping

Over the Krasnodar Region. *11 Edition of the Internacioanl Congresse Geo-Tunis*, *11*(January 2016), 11–20. http://www.ghbook.ir/index.php?name=فرهنگ و رسانه های نوین&option=com_dbook&task=readonline&book_id=13650&page=73&chkhashk=ED9C9491B4&Itemid=218&lang=fa&tmpl=component

Duncan, J., Davies, K. P., Saipai, A., Vainikolo, L., Wales, N., Varea, R., Bruce, E., & Boruff, B. (2022). an Open-Source Mobile Geospatial Platform for Agricultural Landscape Mapping: a Case Study of Wall-To-Wall Farm Systems Mapping in Tonga. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *48*(4/W1-2022), 119–126. https://doi.org/10.5194/isprs-archives-XLVIII-4-W1-2022-119-2022

Duncan, P., Podest, E., Esler, K. J., & Geerts, S. (2023). Mapping Invasive Herbaceous Plant Species with Sentinel-2 Satellite Imagery : Echium plantagineum in a Mediterranean Shrubland as a Case Study Mapping Invasive Herbaceous Plant Species with Sentinel-2 Satellite Imagery : Echium plantagineum in a Mediterra. April. https://doi.org/10.3390/geomatics3020018

Frelat, R., Lopez-Ridaura, S., Giller, K. E., Herrero, M., Douxchamps, S., Djurfeldt, A. A., Erenstein, O., Henderson, B., Kassie, M., Paul, B. K., Rigolot, C., Ritzema, R. S., Rodriguez, D., Van Asten, P. J. A., & Van Wijk, M. T. (2016). Drivers of household food availability in sub-Saharan Africa based on big data from small farms. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(2), 458–463. https://doi.org/10.1073/pnas.1518384112

Ghosh, S., Kumar, D., & Kumari, R. (2022). Cloud-based large-scale data retrieval, mapping, and analysis for land monitoring applications with Google Earth Engine (GEE). *Environmental Challenges*, *9*(May), 100605. https://doi.org/10.1016/j.envc.2022.100605

Gikov, A., Dimitrov, P., Filchev, L., Roumenina, E., & Jelev, G. (2019). Crop type mapping using multi-date imagery from the sentinel-2 satellites. *Comptes Rendus de L'Academie Bulgare Des Sciences*, *72*(6), 787–795. https://doi.org/10.7546/CRABS.2019.06.11

Kapulu, N., Chomba, C., Nkonde, C., Holmes, M., Manda, S., Smith, H. E., Macdiarmid, J. I., & Orfila, C. (2023). Dietary diversity of women from soybean and non-soybean farming households in rural Zambia. *Frontiers in Sustainable Food Systems*, *7*. https://doi.org/10.3389/fsufs.2023.1115801

Kenefick, E. (2004). Thematic Guidelines Sampling Sampling Guidelines for Vulnerability Analysis ODAV (VAM)-WFP, Rome Example-Household Selection Key: Arrow = Random-Walking Direction (spin pencil to determine). December.

Khojely, D. M., Ibrahim, S. E., Sapey, E., & Han, T. (2018). History, current status, and prospects of soybean production and research in sub-Saharan Africa. *Crop Journal*, *6*(3), 226–235. https://doi.org/10.1016/j.cj.2018.03.006

Kolog, J. D., Asem, F. E., & Mensah-Bonsu, A. (2023). The state of food security and its determinants in Ghana: an ordered probit analysis of the household hunger scale and household food insecurity access scale. *Scientific African*, *19*. https://doi.org/10.1016/j.sciaf.2023.e01579

Kuntashula, E., & Mwelwa-Zgambo, L. (2022). Impact of the farmer input support policy on agricultural production diversity and dietary diversity in Zambia. *Food Policy*, *113*(August 2021), 102329. https://doi.org/10.1016/j.foodpol.2022.102329

Li, C., Kandel, M., Anghileri, D., Oloo, F., Kambombe, O., Chibarabada, T. P., Ngongondo, C., Sheffield, J., & Dash, J. (2021). Recent changes in cropland area and productivity indicate unsustainable cropland expansion in Malawi. *Environmental Research Letters*, *16*(8). https://doi.org/10.1088/1748-9326/ac162a

Liang, J., Jin, F., Zhang, X., & Wu, H. (2023). WS4GEE: Enhancing geospatial web services and geoprocessing workflows by integrating the Google Earth Engine. *Environmental Modelling and Software*, *161*(November 2021), 105636. https://doi.org/10.1016/j.envsoft.2023.105636

Lokosang, L. B., Ramroop, S., & Hendriks, S. L. (2011). Establishing a robust technique for monitoring and early warning of food insecurity in post-conflict south Sudan using ordinal logistic regression. *Agrekon*, *50*(4), 101–130. https://doi.org/10.1080/03031853.2011.617902

Mainza, N. O. (2022). Factors Affecting Income Generation Among Smallholder Farmers With Agricultural Cooperative Membership in Zambia. *Journal of Agribusiness and Rural Development*, *64*(2). https://doi.org/10.17306/j.jard.2022.01520

Mango, N., Makate, C., Mapemba, L., & Sopo, M. (2018). The role of crop diversification in improving household food security in central Malawi. *Agriculture and Food Security*, *7*(1), 1–10. https://doi.org/10.1186/s40066-018-0160-x

Mango, N., Zamasiya, B., Makate, C., Nyikahadzoi, K., & Siziba, S. (2014). Factors influencing household food security among smallholder farmers in the Mudzi district of Zimbabwe. *Development Southern Africa*, *31*(4), 625–640. https://doi.org/10.1080/0376835X.2014.911694

McFeeters, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, *17*(7), 1425–1432. https://doi.org/10.1080/01431169608948714

Mubanga, F. C., & Bwalya Umar, B. (2020a). Environmental discounting behaviour of smallholder farmers in Chibombo District, Central Zambia. *Land Use Policy*, *95*(February 2019), 104551. https://doi.org/10.1016/j.landusepol.2020.104551

Mubanga, F. C., & Bwalya Umar, B. (2020b). Environmental discounting behaviour of smallholder farmers in Chibombo District, Central Zambia. *Land Use Policy*, *95*(February), 104551. https://doi.org/10.1016/j.landusepol.2020.104551

Nkembi, L., Herman, D., Mubeteneh, T., & Nkengafac, N. (2021). Analysis of Small Scale Farmers Households Food Security in the Mount Bamboutos Ecosystem. *Journal of Food Security*, *9*(2), 56–61. https://doi.org/10.12691/jfs-9-2-3

Nkonde, C., Audain, K., Kiwanuka-Lubinda, R. N., & Marinda, P. (2021). Effect of agricultural diversification on dietary diversity in rural households with children under 5 years of age in Zambia. *Food Science and Nutrition*, *9*(11), 6274–6285. https://doi.org/10.1002/fsn3.2587

Nuzzo, R. L. (2016). The Box Plots Alternative for Visualizing Quantitative Data. *PM and R*, *8*(3), 268–272. https://doi.org/10.1016/j.pmrj.2016.02.001

Pacheco, P. (2012). Soybean and oil palm expansion in South America. A review of main trends and implications. *Center for International Forestry Research Working Paper 90*, 1–28.

Phalan, B., Bertzky, M., Butchart, S. H. M., Donald, P. F., Scharlemann, J. P. W., Stattersfield, A. J., & Balmford, A. (2013). Crop Expansion and Conservation Priorities in Tropical Countries. *PLoS ONE*, *8*(1). https://doi.org/10.1371/journal.pone.0051759

Phiri, D., Morgenroth, J., & Xu, C. (2019). Four decades of land cover and forest connectivity study in Zambia—An object-based image analysis approach. *International Journal of Applied Earth Observation and Geoinformation*, *79*(December 2018), 97–109. https://doi.org/10.1016/j.jag.2019.03.001

Planque, C., Lucas, R., Punalekar, S., Chognard, S., Hurford, C., Owers, C., Horton, C., Guest, P., King, S., Williams, S., & Bunting, P. (2021). National crop mapping using sentinel-1 time series: A knowledge-based descriptive algorithm. *Remote Sensing*, *13*(5), 1–30. https://doi.org/10.3390/rs13050846

Rao, P., Zhou, W., Bhattarai, N., Srivastava, A. K., Singh, B., Poonia, S., Lobell, D. B., & Jain, M. (2021). Using sentinel-1, sentinel-2, and planet imagery to map crop type of smallholder farms. *Remote Sensing*, *13*(10), 1–13. https://doi.org/10.3390/rs13101870

Savala, C. E. N., Wiredu, A. N., Chikoye, D., & Kyei-Boahen, S. (2022). Prospects and Potential of Bradyrhizobium diazoefficiens Based Bio-Inoculants on Soybean Production in Different Agro-Ecologies of Mozambique. *Frontiers in Sustainable Food Systems*, *6*(July), 1–16. https://doi.org/10.3389/fsufs.2022.908231

Schmitt-Harsh, M., Waldman, K., Estes, L., & Evans, T. (2020). Understanding social and environmental determinants of piecework labor in smallholder agricultural systems. *Applied Geography*, *121*(June), 102243. https://doi.org/10.1016/j.apgeog.2020.102243

Sebatta, C., Wamulume, M., & Mwansakilwa, C. (2014). Determinants of Smallholder Farmers' Access to Agricultural Finance in Zambia. *Journal of Agricultural Science*, *6*(11). https://doi.org/10.5539/jas.v6n11p63

Siamabele, B. (2019). Soya Beans Production in Zambia: Opportunities and Challenges. *American Journal of Agricultural and Biological Sciences*, *14*(1), 55–60. https://doi.org/10.3844/ajabssp.2019.55.60

Siamabele, B. (2021). *The significance of soybean production in the face of changing climates in Africa The significance of soybean production in the face of changing climates in Africa*. https://doi.org/10.1080/23311932.2021.1933745

Simonetti, D., Pimple, U., Langner, A., & Marelli, A. (2021). Pan-tropical Sentinel-2 cloud-free annual composite datasets. *Data in Brief*, *39*, 107488. https://doi.org/10.1016/j.dib.2021.107488

Soofi, E. S. (1990). Effects of collinearity on information about regression coefficients. *Journal of Econometrics*, *43*(3), 255–274. https://doi.org/10.1016/0304-4076(90)90120-I

Soudani, K., Delpierre, N., Berveiller, D., Hmimina, G., Vincent, G., Morfin, A., & Dufrêne, É. (2021). Potential of C-band Synthetic Aperture Radar Sentinel-1 time-series for the monitoring of phenological cycles in a deciduous forest. *International Journal of Applied Earth Observation and Geoinformation*, *104*. https://doi.org/10.1016/j.jag.2021.102505

Sun, L., Chen, J., Guo, S., Deng, X., & Han, Y. (2020). Integration of time series sentinel-1 and sentinel-2 imagery for crop type mapping over oasis agricultural areas. *Remote Sensing*, *12*(1). https://doi.org/10.3390/RS12010158

Talema, T., & Hailu, B. T. (2020). Mapping rice crop using sentinels (1 SAR and 2 MSI) images in tropical area: A case study in Fogera wereda, Ethiopia. *Remote Sensing Applications: Society and Environment*, *18*(November 2019), 100290. https://doi.org/10.1016/j.rsase.2020.100290

Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, *8*(2), 127–150. https://doi.org/10.1016/0034-4257(79)90013-0

Tuni, A., Rentizelas, A., & Chipula, G. (2022). Barriers to commercialise produce for smallholder farmers in Malawi: An interpretive structural modelling approach. *Journal of Rural Studies*, *93*(May), 1–17. https://doi.org/10.1016/j.jrurstud.2022.05.003

Usman, M. A., & Haile, M. G. (2022). Market access, household dietary diversity and food security: Evidence from Eastern Africa. *Food Policy*, *October*, 102374. https://doi.org/10.1016/j.foodpol.2022.102374

Voora, V., Larrea, C., & Bermudez, S. (2020). Global Market Report: Soybeans. *The International Institute for Sustainable Development*, *I*(October), 1–20. https://www.iisd.org/system/files/2020-10/ssi-global-market-report-soybean.pdf

Wambogo, E. A., Ghattas, H., Leonard, K. L., & Sahyoun, N. R. (2018). Validity of the Food Insecurity Experience Scale for use in sub-Saharan Africa and characteristics of food-insecure individuals. *Current Developments in Nutrition*, *2*(9), 6–8. https://doi.org/10.1093/cdn/nzy062

Wang, S., Di Tommaso, S., Deines, J. M., & Lobell, D. B. (2020). Mapping twenty years of corn and soybean across the US Midwest using the Landsat archive. *Scientific Data*, *7*(1), 1–14. https://doi.org/10.1038/s41597-020-00646-4

World Food Program. (2014). *Comprehensive Food Security Survey (CFSS): Yemen*. *April*.

Xie, G., & Niculescu, S. (2022). Mapping Crop Types Using Sentinel-2 Data Machine Learning and Monitoring Crop Phenology with Sentinel-1 Backscatter Time Series in Pays de Brest, Brittany, France. *Remote Sensing*, *14*(18), 1–27. https://doi.org/10.3390/rs14184437

Xu, Y., Yu, L., Zhao, F. R., Cai, X., Zhao, J., Lu, H., & Gong, P. (2018). Tracking annual cropland changes from 1984 to 2016 using time-series Landsat images with a change-detection and post-classification approach: Experiments from three sites in Africa. *Remote Sensing of Environment*, *218*(September), 13–31. https://doi.org/10.1016/j.rse.2018.09.008

You, N., & Dong, J. (2020). Examining earliest identifiable timing of crops using all available Sentinel 1/2 imagery and Google Earth Engine. *ISPRS Journal of Photogrammetry and Remote Sensing*, *161*(December 2019), 109–123. https://doi.org/10.1016/j.isprsjprs.2020.01.001

Yu, J. W., Yoon, Y. W., Baek, W. K., & Jung, H. S. (2021). Forest vertical structure mapping using two-seasonal optic images and lidar dsm acquired from uav platform through random forest, xgboost, and support vector machine approaches. *Remote Sensing*, *13*(21). https://doi.org/10.3390/rs13214282

https://www.zamstats.gov.zm/index.php/publications/category/17-agriculture

7 Appendix A. Survey questionnaire

"Global food trade and local food security: mapping and monitoring of soybean expansion in the Southern region of Africa (using Chibombo District, Zambia as a case study)"

Dear respondent, My name is Chota Bright Nkole final year student at the Faculty of Geo-information and Earth Observation, University of Twente. I am soliciting information on the topic "Global food trade and local food security: mapping and monitoring of soybean expansion in the Southern region of Africa (using Chibombo District, Zambia as a case study)". In partial fulfilment for an MSc degree, I will be glad if you may answer the following questions. Every information provided shall be treated as confidential and used only for academic purposes. Do you voluntarily agree to respond to this survey?

- ◯ Yes
- ◯ no

INTERVIEW INFORMATION

Interview code number

____

Name of the ward

____

Dates of the interview taking place

yyyy-mm-dd

Start time

hh:mm

Name of the interviewer

____

SECTION A: HOUSEHOLD CHARACTERISTICS

I would begin by asking about the head of your household. By household, I mean all the members of the household (including children if any) with whom you share meals; share the same accommodation and other resources like income, land, and equipment.
[Please indicate your answer by selecting the correct answer.

What is the sex of the household head?

- ◯ Male
- ◯ Female

What is the age of the household head?
years

____

What is the level of education

- ◯ Primary
- ◯ Secondary
- ◯ Tertiary
- ◯ No education

Marital status of the Household head?

- ◯ Marriage
- ◯ Single
- ◯ Divorced
- ◯ Widow/widower

What is the household size?
Number of individuals in the house

____

What are the sources of your income?

- ☐ Farming
- ☐ Laborer (e.g. from other peoples farms)
- ☐ remittances
- ☐ Employer (e.g. working as a guard, teacher etc)

What are the sources of food that you consume as the family?

- ☐ Own farm produce
- ☐ Buy from the market
- ☐ Work for food
- ☐ Support from other family members/government, social cash transfer

What is the size of the land do you own as the family?

- ◯ Less than 0.25 ha
- ◯ Between 0.25 and 0.5 ha
- ◯ Between 0.5 and 1 ha
- ◯ Between 1 and 5 ha
- ◯ above 5ha

Do you have assets

- [ ] Tractor
- [ ] Plough
- [ ] Hoes, shovels, rake, sickle etc.
- [ ] Car, lorry, truck, oxen etc.

Do you work?

- ○ Formal employment (doctor , teacher, nurse, etc.
- ○ Informa employment (farmer, guard, hunter, laborer, fisherman etc

Do you hire for the labour?

- ○ Yes
- ○ no

if yes, who do you hire?

- ○ Family members
- ○ Paid laborer's

Do you own livestock?

- ○ Yes
- ○ No

SECTION B: GENERAL CROPS AND SOYBEAN CULTIVATION

Which major crops do you grow?

- [ ] Maize
- [ ] Soybean
- [ ] Groundnuts
- [ ] Millet
- [ ] Cowpeas
- [ ] Cotton
- [ ] Cassava
- [ ] Sweet potatoes
- [ ] Wheat
- [ ] Others

if others, list them?

____

What are the sizes of the area per crop you indicated above?
indicate the name and the size per crop from the above list

____

Which is the most preferred crop ?

____

What is the reason for this preferred crop?

- ◯ Selling for income
- ◯ Home consumption
- ◯ Both

Do you grow soybean?

- ◯ Yes
- ◯ No

If yes , how many years have you been farming soybeans?

____

What is the current hectarage?

____

Have you expanded the soybean area?

- ◯ Yes
- ◯ No

if yes, what was your initial hectarage did you start with?

____

What is the reason for the expansion?

- ◯ Selling for income
- ◯ For home consumption
- ◯ Both

What is the current area (ha) of soybean?

____

Which land use do yo use for expansions?

- ☐ forest, shrubland, virgin land , grassland ets
- ☐ cropland meant for other crops (e.g. the land that was used for sweet potatoes now used for Soybean expansions)
- ☐ Rented land

Do you utilize the soybean that you grow (e.g. for making nutritious food which meets their dietary needs-soy porridge, soymilk etc.)

- ◯ Yes
- ◯ No
- ◯ No knowledge about utilization

Do you do value addition to soybean that you grow (e.g. adding value to soybean to attract more customers, which can boost revenue and profits- for example preprocessed Soybean meal meal that can be sold more than just harvested soybean)).

- ◯ Yes
- ◯ No
- ◯ No knowledge about the value addition

Do you receive support in terms of inputs from the Government or other organizations?

- ◯ Yes
- ◯ No

If yes, select from the list

- ☐ Fertilizer (D-compound, Urea)
- ☐ Maize seed
- ☐ Soybean seed
- ☐ Groundnuts seed
- ☐ Sorghum or Millet seed
- ☐ Others

If others, list them
Type others

____

Do you sell the soybean you produce?

- ◯ Yes
- ◯ No

If yes who buys your soybeans?

- ☐ Government
- ☐ Industries, companies
- ☐ Brief case farmers
- ☐ Others

SECTION C: HOUSEHOLD DIETARY DIVERSITY SCORE

HOUSEHOLD DIETARY DIVERSITY SCORE

To Enumerator: the table below is the dietary diversity score, in which he or she will score a household one (1) if the household ate any food from each of the food groups, otherwise zero (0).

Any , bread, rice noodles, biscuits, cookies, or any other foods made from millet, sorghum, maize, rice, wheat

- ◯ 1
- ◯ 0

Any pumpkin, carrots, squash, or sweet potatoes that are yellow or orange inside?

- ◯ 1
- ◯ 0

Any white potatoes, white yams, manioc, cassava or any other foods made from roots or tubers?

- ◯ 1
- ◯ 0

Any dark, green, leafy vegetables such as cassava leaves, bean leaves, kale, spinach, pepper leaves, taro leaves, and amaranth leaves?

- ◯ 1
- ◯ 0

Any other vegetables?

- ◯ 1
- ◯ 0

Any ripe mangoes, ripe papayas ?

- ◯ 1
- ◯ 0

Any other fruits?

- ◯ 1
- ◯ 0

Any beef, pork, lamb, goat, rabbit wild game, chicken, duck, or other birds, liver, kidney, heart, or other organ meats?

- ◯ 1
- ◯ 0

Any eggs?

- ◯ 1
- ◯ 0

Any fresh or dried fish or shellfish?

- ◯ 1
- ◯ 0

Any foods made from beans, peas, or lentils?

- ◯ 1
- ◯ 0

Any cheese, yoghurt, milk or other milk products?

- ◯ 1
- ◯ 0

Any foods made with oil, fat, or butter?

- ◯ 1
- ◯ 0

Any sugar or honey?

- ◯ 1
- ◯ 0

What is your Household Dietary Diversity Score

- ◯ Low (1-4)
- ◯ Medium(5-8)
- ◯ High (9-12)

Any other foods, such as condiments, coffee, or tea?

- ◯ 1
- ◯ 0

Total score
out of 12

____

SECTION D: HOUSEHOLD FOOD INSECURITY ACCESS SCORE (HFIAS)

To Enumerator: the number of days is the frequency of how many times in a month (e.g.in a month 8 time going to bed hungry then you enter 7)

---

anxiety about food adequacy?

- ◯ Yes
- ◯ No

If Yes, how many times (Frequency)
Enter number of days in the month

____

eating less-preferred foods?

- ◯ Yes
- ◯ No

If yes enter the number of days in the month
Number of days eating less-preferred food in a month

____

eating foods of a limited variety

- ◯ Yes
- ◯ No

If yes, enter number of days
Number of days eating foods of limited variety in a month

____

inability to eat even less-preferred foods

- ◯ Yes
- ◯ No

If yes, enter number of days
Enter number of days of inability to eat even less-preferred foods

____

eating smaller meals than needed

- ◯ Yes
- ◯ No

eating fewer meals in a day

- ◯ Yes
- ◯ No

If yes, enter the number of days
Enter the the number of days of eating fewer meals in a day

____

failing to obtain food of any kind

- ◯ Yes
- ◯ No

If yes , Enter the number of the day

Enter the number of days failing to obtain food of any kind in a month

____

going to bed hungry

- ◯ Yes
- ◯ No

if yes, enter the number of days in a month

enter the number of days of going to bed hungry

____

going the whole day or night without eating anything

- ◯ Yes
- ◯ No

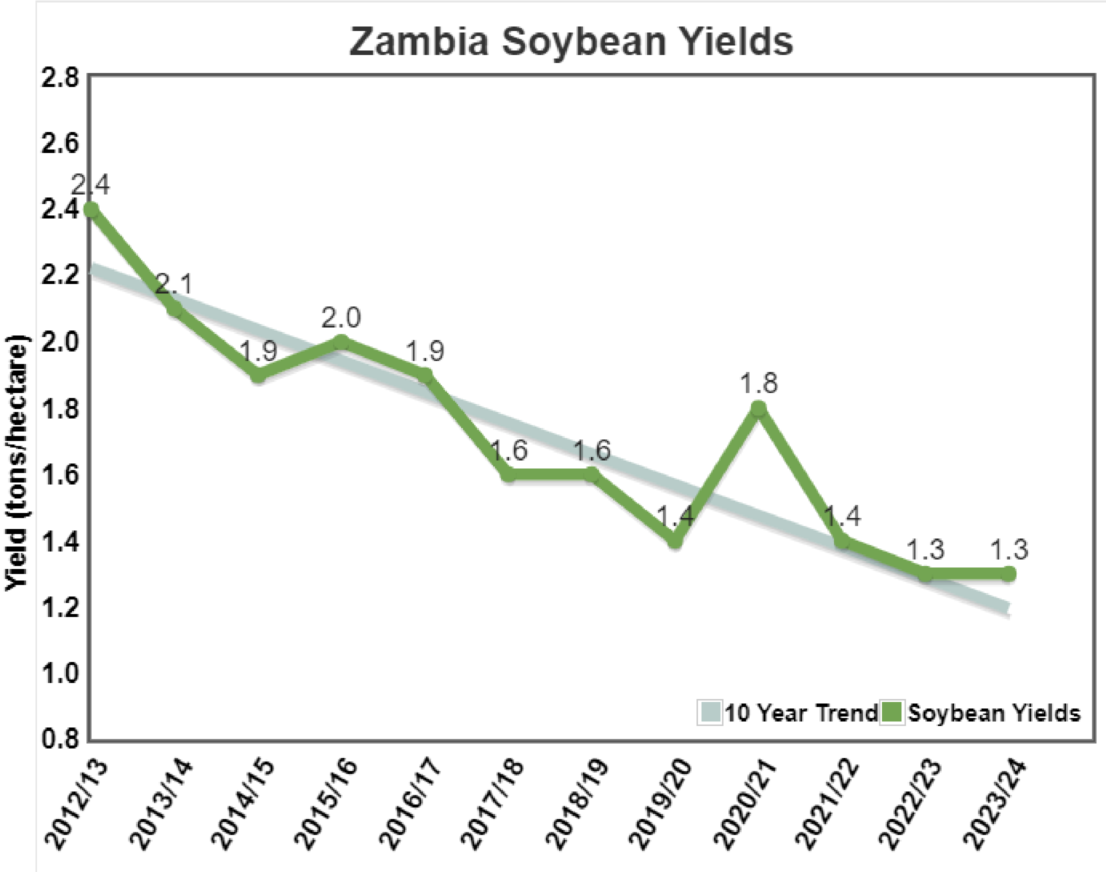
if yes, enter the number of days

enter the number of going whole day or night without eating anything

____

THANK YOU FOR YOUR TIME

8 **Appendix B**. Zambia soybean yields



Source, US Department of Agriculture (2023)