

Article

Delineating Smallholder Maize Farms from Sentinel-1 Coupled with Sentinel-2 Data Using Machine Learning

Zinhle Mashaba-Munghemezulu ^{1,2,*}, George Johannes Chirima ^{1,2} and Cilence Munghemezulu ²

¹ Department of Geography, Geoinformatics and Meteorology, University of Pretoria, Pretoria 0028, South Africa; ChirimaJ@arc.agric.za

² Geoinformation Science Division, Agricultural Research Council Institute for Soil, Climate and Water, Pretoria 0001, South Africa; MunghemezuluC@arc.agric.za

* Correspondence: MashabaZ@arc.agric.za

Abstract: Rural communities rely on smallholder maize farms for subsistence agriculture, the main driver of local economic activity and food security. However, their planted area estimates are unknown in most developing countries. This study explores the use of Sentinel-1 and Sentinel-2 data to map smallholder maize farms. The random forest (RF), support vector (SVM) machine learning algorithms and model stacking (ST) were applied. Results show that the classification of combined Sentinel-1 and Sentinel-2 data improved the RF, SVM and ST algorithms by 24.2%, 8.7%, and 9.1%, respectively, compared to the classification of Sentinel-1 data individually. Similarities in the estimated areas (7001.35 ± 1.2 ha for RF, 7926.03 ± 0.7 ha for SVM and 7099.59 ± 0.8 ha for ST) show that machine learning can estimate smallholder maize areas with high accuracies. The study concludes that the single-date Sentinel-1 data were insufficient to map smallholder maize farms. However, single-date Sentinel-1 combined with Sentinel-2 data were sufficient in mapping smallholder farms. These results can be used to support the generation and validation of national crop statistics, thus contributing to food security.

Keywords: Sentinel-1; Sentinel-2; smallholder; maize; machine learning



Citation: Mashaba-Munghemezulu, Z.; Chirima, G.J.; Munghemezulu, C. Delineating Smallholder Maize Farms from Sentinel-1 Coupled with Sentinel-2 Data Using Machine Learning. *Sustainability* **2021**, *13*, 4728. <https://doi.org/10.3390/su13094728>

Academic Editor: Andrea Colantoni

Received: 23 February 2021

Accepted: 12 April 2021

Published: 23 April 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Maize (*Zea mays L.*) is an essential cereal crop worldwide for food consumption, animal feed, and the production of industrial products such as biofuels [1]. Developed countries consume lower quantities of maize compared to developing countries (Asia, Latin America and Africa), which are reliant on maize [2]. Smallholder farmers account for 80% of the maize produced as a staple crop in Africa [3]. However, global climate forecasts have reported that Africa could be one of the most susceptible regions to the effects of climate change by 2050. This phenomenon will cause growing water shortages and scarcity of suitable land, which will affect the production of cereal crops including maize [4,5]. Smallholder maize farms are important for the livelihoods of rural communities in Africa who depend on agriculture for food security and their local economic activities. These farmers are faced with problems such as inadequate rainfall due to droughts; they often have poor soils and limited irrigation infrastructure, which hinder their maximum productivity [6]. Although these problems prevail in smallholder farms, there is an increasing demand for maize as a consequence of population growth [7]. The disparity between declining maize supply and increasing demand for maize makes it necessary to develop a methodology to map smallholder maize farms and their sizes. Information about the areal extent of smallholder farms will guide the government when dispersing aid to them, inform land-use policies, and provide an indication of the current food security status, especially in vulnerable rural communities. The information provided by this project will enhance initiatives of local governments to provide spatial information regarding agricultural land-use by rural communities, as reliable information is lacking in most developing countries.

The use of remotely sensed data presents an opportunity for mapping smallholder farms and generating spatial information that can support policy implementation and enhance food security planning. Remote sensing technologies are able to collect data over a wide area in near-real time [8]. Additionally, the spatial distribution of crops on other areas within a study location that was not visited can be mapped. However, the use of remote sensing data for mapping smallholder farms has limitations. The coarse spatial resolution of remote sensing products such as the moderate-resolution imaging spectroradiometer (MODIS) and Landsat is not sufficient to map smallholder farmland plots because of their small size of ± 2 ha. Additionally, Landsat 8 has a revisit cycle of 16 days, which is insufficient to capture phenological changes for smallholder farms. Other remote sensing products such as Worldview, PlanetScope, RapidEye, and Satellite Pour l'Observation de la Terre (SPOT) have the required spatial resolution but are not freely available, and have a limited spatial coverage [9]. Hence, there is a need to explore the use of Sentinel-1 and Sentinel-2 data, which are freely available and have an improved spectral and spatial resolution.

The Sentinel-1 and Sentinel-2 sensors were launched for different applications amongst others, monitoring land-use/land-cover change and agricultural applications [10]. These sensors have a shorter revisit time of 10–12 days and a spatial resolution of 10–60 m [11]. Sentinel-2 is an optical sensor, which captures changes in land cover and provides a means to estimate crop area. However, the optical data from Sentinel-2 are susceptible to cloud cover or rainy weather, which limits the data availability during the cropping season [12]. Radar imagery from Sentinel-1 overcomes the above shortfall; data are unobstructed by clouds or weather. These data have not been explored extensively for agricultural applications in comparison to optical data because of their complex data structure [13].

The combined use of both Sentinel-1 and Sentinel-2 has the advantage of capturing both the spectral and textural information; this improves classification results, according to Cai et al. [14]. Dobson [15] also observed that other Synthetic Aperture Radar (SAR) data such as ERS-1 and JERS-1 are also sensitive to the structural properties, soil moisture and above-ground biomass of vegetation. Studies combining both Sentinel-1 and Sentinel-2, such as that of Van Tricht et al. [16], have found overall accuracies (OA) between 75 and 82% when mapping maize and other land-cover classes with the application of Random Forest (RF) classification with Sentinel-1 and Sentinel-2 data. Sonobe et al. [17] used a kernel-based extreme learning machine to map maize and other crop types with Sentinel-1 and Sentinel-2 data. Their study achieved an overall classification of 96.8%. To our knowledge, limited studies have explored the potential offered by combining radar and optical data to address smallholder crop classification/mapping in a rural setting.

We examined the utility of Sentinel-1 to mapping smallholder areas under maize. We determined the outcome of integrating optical bands and vegetation indices derived from Sentinel-2 on the Sentinel-1 polarizations through a series of classification experiments for mapping maize areas. The RF algorithm, support vector machine (SVM) algorithm and model stack (ST) are applied to each experiment. These machine learning algorithms are selected specifically because they have a superior discrimination capacity between different classes, suitable for noisy data and can be applied to limited samples [18,19]. These distinguishing characteristics of the selected models have the potential of resolving issues with mapping fragmented inhomogeneous smallholder farms. Thereby, we achieved the overall aim of the study in developing a framework to enhance the delineation of smallholder maize farms using Sentinel-1, Sentinel-2 and vegetation indices.

2. Materials and Methods

2.1. Study Area

The field data were collected from the Makhuduthamaga district in Limpopo, South Africa (Figure 1). This area experiences rainfall during the warmer months of October to March and the mean annual rainfall is 536 mm. The fields have an average elevation of 1333 m above mean sea level. The temperatures can drop to 7 °C in winter but can be as

high as 35 °C in summer according to the records from the automatic weather stations of the Agricultural Research Council. This area was selected as a case study because most of the rural population are smallholder maize farmers; they farm primarily for subsistence and partially for selling in local markets [20]. Specific regions of interest (ROI) were delineated for investigation based on the locations of the smallholder maize farms. The ROI was obtained from the local government department of agriculture (DAFF), where they were developed through survey campaigns. The ROI was used to generate an improved estimate of the area covered by smallholder farms by eliminating built-up areas, which can host households with backyard maize gardens leading to an overestimation of the planted areas. These households consume their maize before harvest-time.

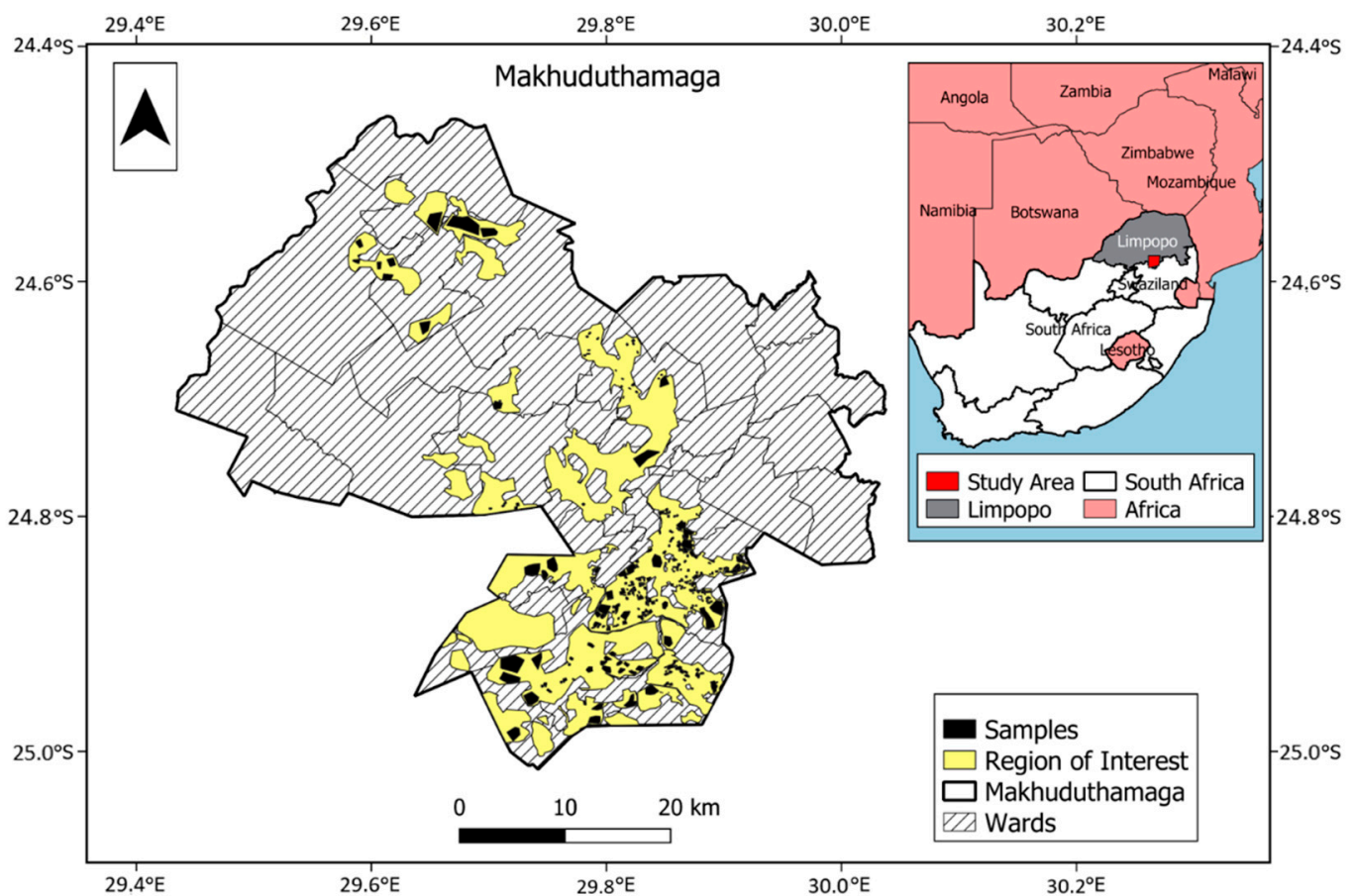


Figure 1. Location of Makhuduthamaga study area within Limpopo province, South Africa.

2.2. Field Data Collection

Field surveys for the collection of training and validation data for different landcover types within the ROI occurred from 18 to 21 February 2019. This period was selected because maize had the maximum green biomass at this time and could be discriminated more clearly in comparison to other land-cover types [21]. A handheld Garmin Global Positioning System (GPS) device was used to collect waypoints of different land-cover classes, applying a purposive sampling approach. The classes considered were maize (19.72%), bare land (50.01%), vegetation (30.23%) and water (0.0%), which are the dominant classes in the study area. The bare land, vegetation and water classes were amalgamated to form the non-maize areas and the maize areas were used as well. This approach of using only two classes of (1) maize and (2) non-maize areas reduces the classification errors from incorporating different land-cover classes individually. For example, there were fewer pixels for water in the study area in comparison to bare land and vegetated areas; using

this as a separate class has the potential of introducing errors depending on the sensitivity of the classifier. Ground-based validation samples for 18 smallholder maize farms were collected using a GPS. The samples were not used as training data for classification.

2.3. Sentinel-1 Data Acquisition and Pre-Processing

Sentinel-1 Level-1 ground range detected (GRD) data described in Table 1 were acquired from the Copernicus Open Access Hub. The interferometric wide (IW) image for 20 February 2019 was used; this consisted of the vertical transmit and vertical receive (VV) and vertical transmit and horizontal receive (VH) polarized backscatter values (in decibels) in a 10 m spatial resolution. Pre-processing of the radar images was done using the Sentinel application platform (SNAP). The orbit file was applied to update the orbit state vectors in the metadata file. Then, radiometric calibration was performed to convert the intensity values into sigma nought values. Speckle filtering was implemented to remove the granular noise caused by the interference of waves reflected from many scatterers. The Lee filter was applied at a 7×7 window size as it was found to be superior in preserving the edges, linear features, point target and texture [22]. Range Doppler terrain correction was done to correct for geometric distortions caused by topography such as foreshortening and shadows; the Shuttle Radar Topography Mission (SRTM) 3-sec Digital Elevation Model (DEM) was used for this purpose [23]. The backscatter values were converted into decibels, then the VH and VV polarizations were used to generate the VV/VH ratio.

Table 1. Specifications of the Sentinel-1 and Sentinel-2 MSI data used in this study.

Spectral Band/Polarization	Central Wavelength (nm)	Bandwidth (nm)	Spatial Resolution (m)
Sentinel-1			
Vertical transmit and vertical receive (VV)	55,465,763	-	10
Vertical transmit and horizontal receive (VH)	55,465,763	-	10
Sentinel-2 MSI			
2-Blue	490	65	10
3-Green	560	35	10
4-Red	665	30	10
5-Vegetation Red Edge (RE1)	705	15	20
6-Vegetation Red Edge (RE2)	740	15	20
7-Vegetation Red Edge (RE3)	783	20	20
8-Near-Infrared (NIR)	842	115	10
8a-Vegetation Red Edge (RE4)	865	20	20
11-Short-wave Infrared (SWIR1)	1610	90	20
12-Short-wave Infrared (SWIR2)	2190	180	20

2.4. Sentinel-2 Data Acquisition and Pre-Processing

The Sentinel-2 Level-1C image for 26 February 2019 was acquired from the Copernicus Open Access Hub. The Sentinel-2 images were pre-processed using the Sen2Cor plugin in SNAP to convert them from the top of atmosphere reflectance units to the bottom of atmosphere reflectance [24]. The bands which were used are summarized in Table 1. The SWIR and vegetation red edge bands were rescaled to 10 m resolution. The indices depicted in Table 2 were derived. These indices are necessary to be investigated for mapping smallholder farms because they cover a broad part of the electromagnetic spectrum (NIR, red and green) in comparison to only using the normalized difference index (NDVI). Additionally, they are sensitive to changes in soil background; they enhance the green vegetation signal, reduce the saturation effect of NDVI and are sensitive to chlorophyll content [25–34].

Table 2. Vegetation indices computed from Sentinel-2 imagery.

Vegetation Index	Equation	Justification	Reference
DVI	$DVI = NIR - Red$	Distinguishes between maize and soil.	[26]
GNDVI	$GNDVI = (NIR - Green) / (NIR + Green)$	More sensitive to chlorophyll concentration than NDVI.	[31]
IPVI	$IPVI = NIR / (NIR + Red)$	Similar to NDVI, but it is computationally faster.	[28]
MSR	$MSR = \left(\frac{NIR}{Red} - 1 \right) / \left(\left(\frac{NIR}{Red} \right)^{1/2} + 1 \right)$	Minimizes the effects of variable soil reflectance.	[30]
MTVI1	$MTVI1 = 1.2 \times [1.2 \times (NIR - Green) - 2.5 \times (Red - Green)]$	Predicting maize green LAI (leaf area index).	[34]
MTVI2	$MTVI2 = \frac{1.5 \times [1.2 \times (NIR - Green) - 2.5 \times (Red - Green)]}{\sqrt{(2 \times NIR + 1)^2 - (6 \times NIR - 5 \times \sqrt{Red})} - 0.5}$	Better predictor of maize green LAI than MTVI1, and it accounts for soil background.	[34]
NDVI	$NDVI = \frac{NIR - Red}{NIR + Red}$	Sensitive to maize greenness. However, it can saturate in dense vegetation when LAI becomes very high.	[26]
OSAVI	$OSAVI = \frac{NIR - Red}{NIR + Red + 0.16}$	Eliminates the effect of the soil background.	[32]
RDVI	$RDVI = \frac{NIR - Red}{\sqrt{NIR + Red}}$	Detects maize and is not sensitive to the effects of soil and sun viewing geometry.	[29]
SAVI	$SAVI = \frac{((1+L) \times (NIR - Red))}{NIR + Red + L}$	The SAVI index is similar to NDVI, but it reduces the influence of soil.	[27]
SR	$SR = \frac{NIR}{Red}$	Detects healthy maize. However, it can saturate in densely vegetated maize plots when LAI becomes very high.	[25]
TVI	$TVI = 0.5 \times [120 \times (NIR - Green) - 200 \times (Red - Green)]$	Detects green maize biomass and chlorophyll.	[33]

Note: DVI: difference vegetation index; GNDVI: green normalized difference vegetation index; IPVI: infrared percentage vegetation index; MSR: modified simple ratio; MTVI1: modified triangular vegetation index; MTVI2: modified triangular vegetation index—modified; NDVI: normalized difference vegetation index; OSAVI: optimized soil-adjusted vegetation index; RDVI: renormalized difference vegetation index; SAVI: soil-adjusted vegetation index; SR: simple ratio; TVI: triangular vegetation index.

2.5. Classification Algorithms

Three different approaches were applied for mapping the smallholder farms, namely, RF, SVM and ST. The RF algorithm is a non-parametric decision tree ensemble classifier [35]. This classifier consists of a large number of classification and regression trees (CART), where each pixel is classified using a majority voting system. The RF algorithm trains each tree using an independently drawn subset of the original data using bootstrapping or bagging, and determines the number of features to be used at each node through an evaluation of a random vector [35]. One tuning parameter was defined for RF, the number of trees to grow (ntree), and the rest of the parameters are set to default values. In this study, the ntree was 150; this minimized the Out of Bag error, similar to Rodriguez-Galiano et al. [36]. The RF algorithm was selected because it can handle high dimensional data, is less sensitive to over-fitting and makes no distribution assumptions [18,37,38].

The SVM algorithm is also a non-parametric supervised learning classifier. The SVM uses the kernel function to transform training data into a high dimensional feature space, and to identify the optimal hyperplane that maximizes the distance between the separating hyperplane and the nearest sampling points [39–41]. The radial basis kernel was applied for SVM because of its good performance in previous studies [42,43]. The regularization parameter, gamma value and kernel coefficient had to be defined for the classifier. In this study, the regularization parameter was 100, the gamma value was 0.01 and kernel coeffi-

cient was 0, similar to Kumar et al. [44]. The SVM algorithm was selected because it does not make assumptions of the probability distribution and is not sensitive to training sample size [40]. A grid-search method was used to find these optimum turning parameters for both SVM and RF.

Model stacking was applied; it collates the predictions generated by different machine learning algorithms and uses them to generate a second-level learning classifier [45]. In this study, the RF and SVM classifier were stacked, and the Logistic Classifier was used to combine the results. This ensemble model was applied because it has the ability to increase the predictive capacity of the two classifiers instead of using them independently [45].

Although RF has a variable importance measure, the permutation feature importance measurement was applied in this study to determine the importance of the predictors in each experiment, since previous studies have shown that RF variable importance has variations in ranking predictors as different iterations are performed [46]. The permutation feature importance allows different trained models (RF, SVM and ST) to assess feature importance. The algorithm computes reference scores s for the selected model on experimental datasets D . This reference score is the overall accuracy of the classifier. The features j in the datasets D are randomly shuffled to generate a corrupted version of the data $\tilde{D}_{k,j}$. The scores $s_{k,j}$ are computed on the corrupted datasets $\tilde{D}_{k,j}$. The feature importance i_j is then computed for feature f_j according to Equation (1)

$$i_j = s - \frac{1}{K} \sum_{k=1}^k s_{k,j}. \quad (1)$$

2.6. Experimental Design

These samples were randomly separated into training (80% of the data) and testing (20% of the data) [47]. The training data were used for classification, whereas the testing data were used to evaluate the models. The vegetation indices in Table 2 were derived for use during classification. Then, classification experiments depicted in Table 3 were set for the classification algorithms based on different combinations (data configurations). These experimental set-ups were adopted to investigate the best approach for mapping smallholder maize with Sentinel-1 and Sentinel-2 data.

Table 3. Combinations (data configurations) for the four experiments.

Experiment Number	Combinations	Description
1	VH, VV, VV/VH	Sentinel-1 polarization
2	VH, VV, VV/VH, DVI, GNDVI, IPVI, MSR, MTVI1, MTVI2, NDVI, OSAVI, RDVI, SAVI, SR, TVI	Sentinel-1 polarization and vegetation indices
3	VH, VV, VV/VH, DVI, GNDVI, IPVI, MSR, MTVI1, MTVI2, NDVI, OSAVI, RDVI, SAVI, SR, TVI, IPVI, 2, 3, 4, 5, 6 7, 8, 8a, 11, 12	Sentinel-1 polarization, vegetation indices and Sentinel-2 bands
4	VH, VV, VV/VH, 2, 3, 4, 5, 6 7, 8, 8a, 11, 12	Sentinel-1 polarization, and Sentinel-2 bands

2.7. Classification Model Evaluation and Planted Maize Area Estimation

Model evaluation was done to select the ideal model for estimating the maize areas. The matrices used were the OA, kappa coefficient of agreement (\hat{k}), cross-validation, precision, recall and F1-Score. The OA is the total classification accuracy and values close to 1 indicate that a classification is accurate; this is computed according to Equation (2). The OA was adjusted using the procedure of Olofsson et al. [48] to account for classification errors. The \hat{k} is calculated according to Equation 3 where k is the land-cover classes in the confusion matrix, x_{i+} and x_{+j} represent the marginal total for row i and column j . x_{ii} represents the number of observations in the row i and column i and N represents the total number of samples. \hat{k} values > 0.8 represent a strong agreement between the classification map and the ground reference data. \hat{k} values between 0.4 and 0.8 represent

moderate agreement and \hat{k} values < 0.4 represent poor agreement [49]. The equations for both matrices are given as:

$$\text{Overall accuracy} = \frac{\sum_{i=1}^k x_{ii}}{N}, \quad (2)$$

$$\hat{k} = \frac{N \sum_{i=1}^k x_{ii} - \sum_{i=1}^k (x_{i+} \times x_{+j})}{N^2 - \sum_{i=1}^k (x_{i+} \times x_{+j})}. \quad (3)$$

The K-fold cross-validation method was then applied [50]. This method divides the training data randomly into K-folds or subsets (in this study a standard value of 10 was used), where one of the subsets is used as a test data set and the other K-1 is used as a training data set used to fit the model. This process is repeated i times, and the calculated average accuracy is computed for the testing data. The accuracy statistic was used during cross-validation, where values close to 1 indicate a high probability that a sample is correctly classified. The standard deviation of each accuracy value is also computed in each iteration, and the average standard deviation is indicated using a $+/-$ attached to the cross-validation accuracy. The precision, recall and F1-Score were computed to determine the rate at which the pixels were correctly classified. The classifier performs well if the precision, recall and F1-Score are close to 1 [51].

Classification confidence was evaluated using McNemar's test to compare each of the models together [52]. We tested the hypothesis that the two models perform the same. When the Chi-squared values are less or equal to 3.84, the models have the same error at a 95% confidence level. However, one model is superior if the Chi-squared values are greater than 3.84.

The areas derived from the classification map were adjusted to account for classification error, and the 95% confidence interval was computed to compare the three models [48]. These areas were compared to the areas derived from 18 maize farms measured during fieldwork to get an indication of how accurately the models estimate maize-planted areas using a regression equation. The p-value (p) and Pearson correlation coefficient (R) are used to evaluate the accuracy.

3. Results

3.1. Classification Model Evaluation

The performances of the three algorithms applied in this study are presented in Table 4. The experiment with the lowest accuracies was experiment 1, containing the Sentinel-1 polarizations independently. This experiment had an accuracy of between 0.68–0.85 and a cross-validation score of between 0.65–0.69 for the three algorithms. Furthermore, the precision (0.65–0.69), recall (0.60–0.71) and F1-Score (0.64) for this experiment were considerably lower than all the other experiments. The kappa values also indicate moderate agreement between models and the reference data. However, there was a notable increase in accuracy by adding vegetation indices to the Sentinel-1 polarizations. The vegetation indices increased the accuracies by 24.2% for RF, 8.7% for SVM and 9.1% for ST. Although there was a reasonable improvement in model performance (precision of 0.925–0.929, recall of 0.926–0.930 and F1-Score of 0.925–0.930) in this experiment, adding Sentinel-2 bands improved the performance further by 5.9% for RF, 5.7% for SVM and 5.8% for ST in experiment 3. The best-performing experiment for all algorithms was experiment 4 with Sentinel-1 polarization and Sentinel-2 bands. This experiment had the highest accuracy (0.99) and was the most accurate (cross-validation: 0.91–0.92, precision: 0.99, recall: 0.99, and F1-Score: 0.99). McNemar's test (Table 5) confirmed that all three algorithms had a different performance in experiments 1–3. However, the performances of the algorithms were similar for the ST-RF combination but different for the ST-SVM combination in experiment 4.

Table 4. The model performance statistics for the three classifications (RF–random forest, SVM–support vector machine, ST–model stack) algorithms in different experimental (Exp) setups.

Exp	Algorithm	Overall Accuracy	Cross-Validation	Precision	Recall	F1-Score	Kappa
1	RF	0.679	0.647 +/- 0.131	0.652	0.660	0.637	0.509
	SVM	0.845	0.688 +/- 0.127	0.693	0.706	0.640	0.526
	ST	0.841	0.689 +/- 0.128	0.674	0.703	0.637	0.523
2	RF	0.921	0.869 +/- 0.118	0.926	0.927	0.926	0.885
	SVM	0.932	0.873 +/- 0.112	0.925	0.926	0.925	0.884
	ST	0.932	0.870 +/- 0.109	0.929	0.930	0.930	0.890
3	RF	0.980	0.903 +/- 0.127	0.983	0.983	0.983	0.972
	SVM	0.989	0.883 +/- 0.106	0.991	0.991	0.991	0.986
	ST	0.990	0.899 +/- 0.137	0.991	0.991	0.991	0.986
4	RF	0.987	0.907 +/- 0.132	0.989	0.989	0.989	0.982
	SVM	0.991	0.914 +/- 0.082	0.992	0.992	0.992	0.988
	ST	0.991	0.921 +/- 0.112	0.991	0.991	0.991	0.986

Table 5. McNemar’s test results for the ST–RF and ST–SVM combinations for experiments 1–4.

Combination	Chi-Squared	p-Value
ST1–RF1	4396.2	0
ST1–SVM1	430	1.7×10^{-95}
ST2–RF2	120	6.3×10^{-28}
ST2–SVM2	516.5	2.4×10^{-114}
ST3–RF3	6.3	0.0002
ST3–SVM3	34.5	4.2×10^{-9}
ST4–RF4	0.05	0.83
ST4–SVM4	9.3	0.0002

3.2. Variable Importance

The variable importance was determined for the experiments in Table 3 using the permutation feature importance algorithm [46]. The experiments (Figure 2) varied in terms of the most important predictors depending on the input data. In experiment 1, the VH polarization had the highest importance; however, when integrating other predictors (e.g., experiments 3 and 4), the VV polarization had a higher importance over the other polarizations. The DVI outperformed all the other vegetation indices, followed by GNDVI in experiment 2. The most important bands in experiments 3 and 4 were the blue, red-edge and short-wave infrared (SWIR) bands. Additionally, the Sentinel-2 spectral bands took the highest priority in terms of importance in comparison to the Sentinel-1 polarizations.

3.3. Mapping and Area Estimates for Maize

The 95% confidence interval was computed for the maize and non-maize areas within the study area. There was a relatively small variation between the total areas classified by the three algorithms for maize in Table 6. The RF algorithm had a discrepancy of 6% when compared to SVM, and 0.7% when compared to ST for the maize-planted areas. The ST algorithm had a variation of 5.5% in comparison to SVM. The areas classified as planted with maize had a lower error (0.7–1.2 ha) in comparison to the other areas which were not maize (1.2–1.88 ha) based on the 95% confidence interval. The RF algorithm had the lowest accuracy of ± 1.2 ha when estimating maize areas, and SVM had the highest accuracy of ± 0.7 ha.

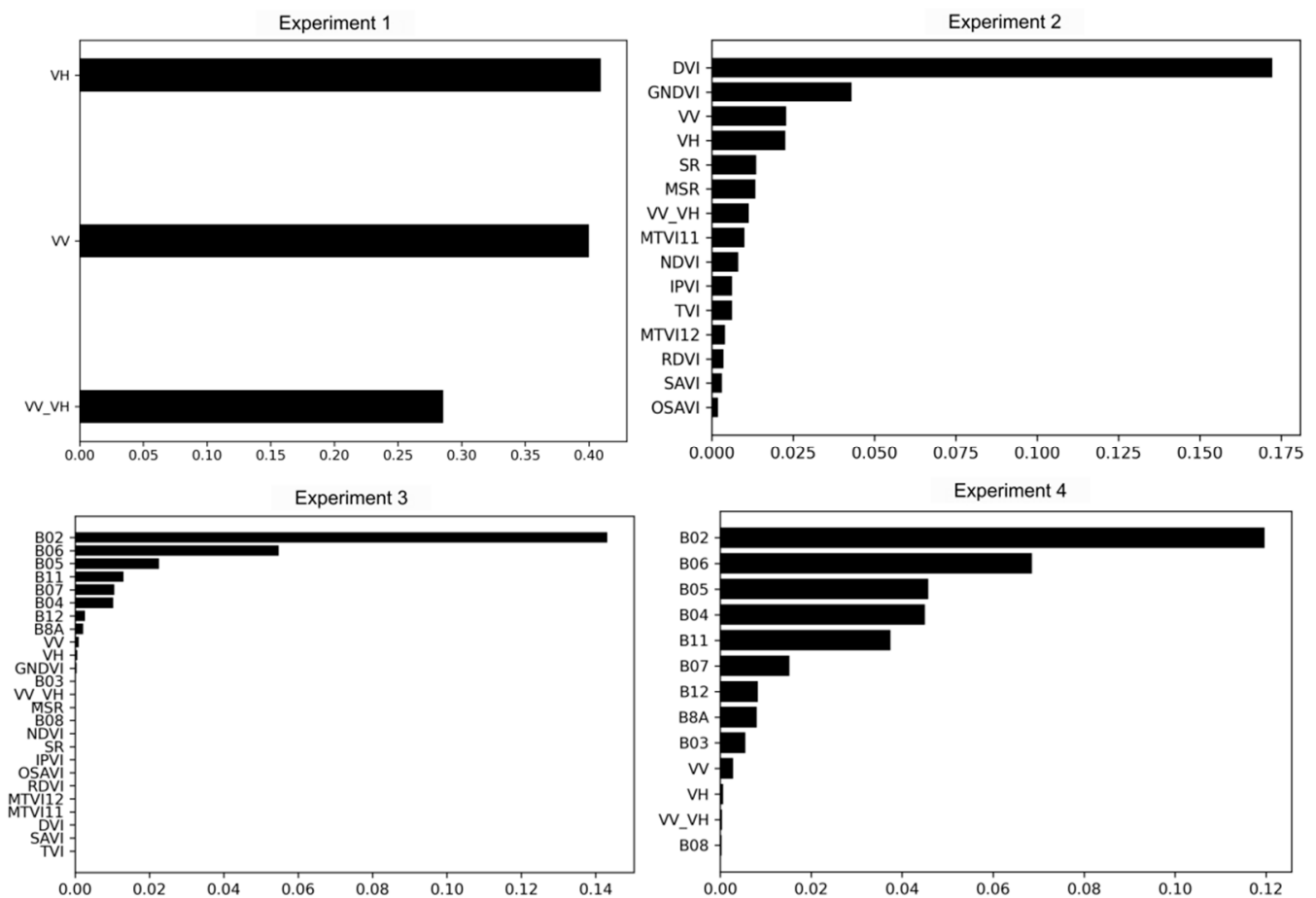


Figure 2. Variable importance plot for the four experiments.

Table 6. Estimated areas based on experiment 4 generated by the three classifiers for maize-planted areas and non-maize areas.

Algorithm	Land-Cover	Total Area (ha)	95% Confidence Interval (ha)
RF	Maize	7001.35	1.236
	Non-Maize	33,496.05	1.884
SVM	Maize	7926.03	0.735
	Non-Maize	32,571.37	1.242
ST	Maize	7099.59	0.819
	Non-Maize	33,397.81	1.202

The classified areas for 18 smallholder maize farms were related to the field measured area at the same farms in Figure 3. There was a positive relationship, which was significant at a 95% confidence interval ($p < 0.05$) between the classified areas and field measured areas. The correlation coefficients obtained by the RF, ST and SVM algorithms are 0.51, 0.78 and 0.84, respectively, indicating higher agreement with the field measurements.

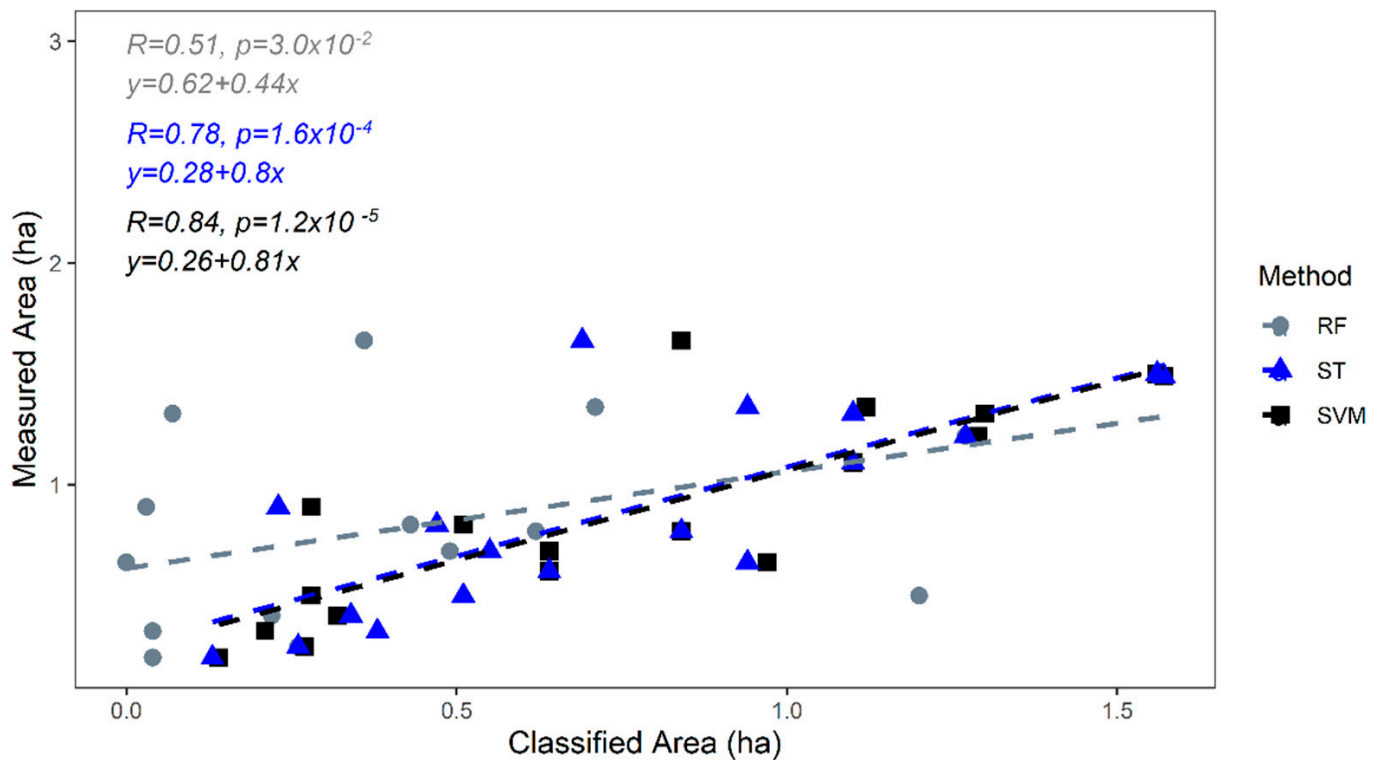


Figure 3. Linear regression models for the field measured areas (y) compared to the classified areas (x) for the best-performing experiment (experiment 4).

The three algorithms were used to generate the classification maps in Figure 4b–d depicting the spatial patterns of the two classes considered within the ROI. These maps compared well with the true color composite satellite image in Figure 4a for Sentinel-2. The classification maps generated by SVM, RF and ST were similar. The maize-planted areas were concentrated in the southern part of the Makhudutamaga district. The crop maps derived in this study are fundamental for crop forecasting and crop yield estimation at the end of the season. Changes induced by natural phenomena, such as climate variability and their effects on crop production, can be understood with the use of crop maps.

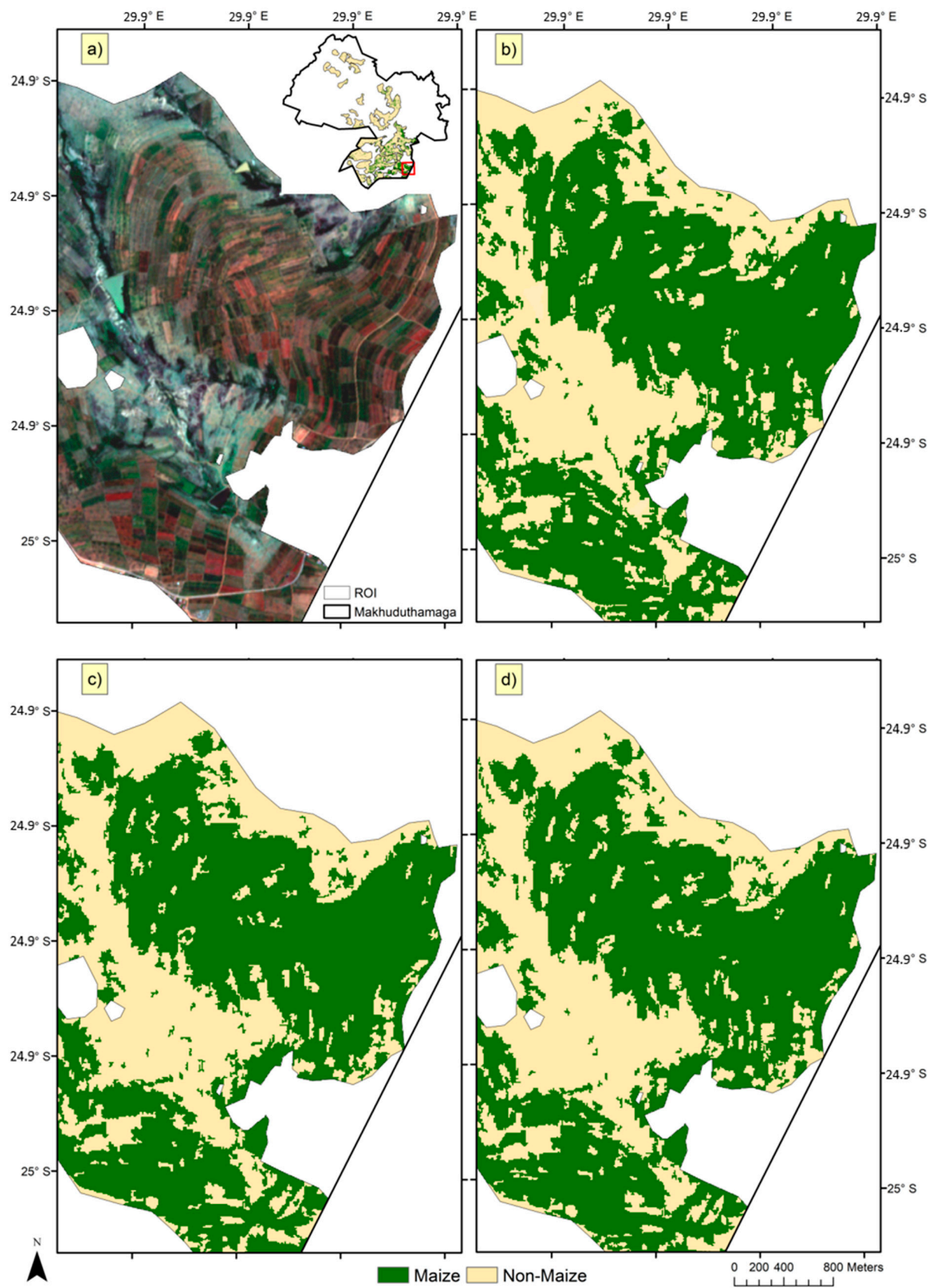


Figure 4. Classification maps for the optimal performing models in experiment 4, where (a) is the true color composite, (b) is RF, (c) is SVM and (d) is ST.

4. Discussion

This study assessed the applicability of Sentinel-1, Sentinel-2 and derived vegetation indices for mapping smallholder maize in Makhudutamaga, Limpopo Province. Classification experiments were set to evaluate the performance of three machine learning algorithms. The variable importance measures were employed to investigate which predictors had the most influence in each experiment. The best performing algorithms were then used for estimating and mapping the maize-planted areas. Findings suggest that integrating Sentinel-1 and Sentinel-2 is ideal for mapping smallholder maize farms with the application of machine learning algorithms.

Contrary to our expectations, the use of single-date Sentinel-1 radar data was not effective for mapping smallholder maize farms. The data combination consisting of Sentinel-1 polarizations exclusively had a low OA ranging from 67.9% to 84.5%, with RF being the worst performing classifier. These results are similar to those of Abubakar et al. [53], who observed an OA of 78.9% when mapping smallholder maize using Sentinel-1 data by applying SVM. However, Useya and Chen [54] reported an OA of 46% with RF and 40% with K-means classification when mapping smallholder maize farms and other crops with Sentinel-1 single-date data. The poor performance of the Sentinel-1 C-band data could be because of its shorter wavelength, which decreases canopy penetration in comparison to L-band SAR, which has a longer wavelength [55,56]. The inconsistencies in the planting pattern in the smallholder farms, such as a lack of equal row spacing, differences in the plant densities, leaf area index and crop heights in the study area, detract from the performance of the Sentinel-1 data because, according to Inoue et al. [57], C-band data are sensitive to changes in biomass.

The integration of Sentinel-1, Sentinel-2 and vegetation indices were ideal for detecting smallholder maize farms, similar to previous studies in comparison to using Sentinel-1 data independently. Experiments 2, 3 and 4 show a clear increase in performance measures, in both OA and cross-validation scores. These values are more consistent and similar to each other, indicating the positive impact of radar-optical fusion on classification accuracy. Other studies such as that of Van Tricht et al. [16] achieved OAs between 75 and 82% when mapping maize and other land-cover classes with the application of Sentinel-1 and Sentinel-2 data. Abubakar et al. [53] achieved an OA of 97% when mapping smallholder maize with vegetation indices, Sentinel-1 and Sentinel-2 data. The high accuracies attained in this current study are attributed to the use of ideal locations of the electromagnetic spectrum such as the red-edge and SWIR. Furthermore, the vegetation indices applied in the current study reduce background effects (soils and other classes such as buildings), thereby enhancing the detection of crops and vegetation classes [25–34].

The differences in performance of the SVM, RF and ST algorithms were expected. For example, Ouzemou et al. [58] reported different OAs of 89.3%, 85.3% and 57.2% for RF, SVM and Spectral Angle Mapper (SAM) for crop type mapping with Landsat 8 data. Sonobe et al. [59] found that SVM (OA of 89.1%) had a superior performance than RF (OA of 87.8%) and the Classification and regression tree (CART) (OA of 81.2%) algorithms for classifying crops with TerraSAR-X data. These differences can be induced by various factors. In this study, the first experiment had the lowest accuracies; notably, RF had a low performance. This is because RF has been shown to be highly sensitive to small number of training input data in previous studies, in comparison to SVM and ST [60,61]. All three algorithms had high accuracies in the four experiments, possibly because the ROI used for training focused on maize-planted areas. This approach reduced the effects of using multiple land-cover classes individually which has a potential to lower the classification accuracy.

The variable importance results indicating the superiority of the VV polarization, DVI, GNDVI, blue band, red-edge and SWIR bands for mapping maize were expected. Forkuor et al. [62] found that the VV band was superior to the VH band derived from TerraSAR-X for crop mapping applications. Deschamps et al. [63] used Sentinel-1 data for crop classification and observed that the VV band was important for crop classification. However, other studies, for example Inglada, et al. [64] and Arias et al. [65], have reported that the VH band

is more important than the VV bands for mapping crops because it captures the volume scattering from the crop canopy structure [66]. These results suggest that it is important to evaluate the polarizations based on the locality where they are applied. The finding that DVI and GNDVI are the most important indices, when using radar data and vegetation indices for crop classification, highlights the importance of evaluating different indices instead of relying on the commonly used NDVI index. The blue band, red-edge and SWIR bands have proven to be important in previous studies [38,67,68]. These bands capture the biochemical properties, water content and residue cover of different crop types that improves their detection [69]. In experiment 2, the OSAVI index was the least important variable. However, this seems to change in experiment 3, where this index ranked higher than RDVI, MTV12, MTV11, DVI, SAVI and TVI. This may be due to the correlation of these bands with the raw Sentinel-2 bands in experiment 3, while the indices in experiment 2 have a lower correlation between them.

The RF and ST algorithms had a relatively small difference of 0.7% when estimating the total planted maize area class, while the SVM algorithm seems to have overestimated the planted maize area by approximately 6% compared to the results from other algorithms. Even though SVM had a higher correlation coefficient than the RF and ST algorithms, we could not conclude that the SVM was the better estimator since the validation samples are relatively small. More validation data are required to provide more information on the performance of each algorithm in relation to ground-measured areas. However, since all algorithms have similar positive values of correlation coefficients, we can conclude that these algorithms can be used to estimate smallholder maize farmed areas. Unfortunately, official agricultural statistics such as production areas are not available in our study area, and could have been used to validate these observations.

The findings of this study are applicable to the Sustainable Development Goals (SDG), specifically, SDG number 2 (Zero Hunger), target 2.4 and indicator 2.4.1, which concern mitigating factors that affect agricultural production, ensuring sustainable agriculture and increasing the proportion of agricultural area under production [70]. The agricultural production area is of great importance, as it informs local government and related stakeholders about agricultural activities and provides means by which production can be forecasted. The production area is one of the important indicators of food insecurity, especially in developing countries such as South Africa. Thus, this study contributes towards this SDG by using remote sensing data to accurately map production areas for smallholder maize farms. The spatial information generated can be used by local government to assist smallholder farms and policy implementation [70].

The limitations of this study were that a limited number of sample points were collected during fieldwork due to the undulating nature of the terrain, high cost to conduct the fieldwork and prominent mountainous areas, which were not accessible for data collection. This small sample size affects the statistical robustness of results [71]. Secondly, the poor farm management practices of smallholder farmers such as weeds and patches of grass growing in some of the farms affect the spectral signature of maize and decrease the accuracy at which they can be detected with remotely sensed imagery. Thirdly, the use of red-edge indices, which have demonstrated some potential in improving the detection of vegetation in previous studies, should be explored [72,73].

5. Conclusions

The overall aim of the study was to develop a framework to enhance the delineation of smallholder maize areas using single-date Sentinel-1, Sentinel-2 and derived vegetation indices. The results showed that single-date Sentinel-1 on its own was not sufficient in mapping planted maize fields. When Sentinel-2 data were integrated with Sentinel-1 data, an improvement of 24.2%, 8.7% and 9.1% for RF, SVM and ST algorithms, respectively, were observed. Machine learning proved to have a high capacity to estimate smallholder maize-planted areas (7001.35 ± 1.2 ha for RF, 7926.03 ± 0.7 ha for SVM and 7099.59 ± 0.8 ha for ST). The framework used in this study can be applied when evaluating different algorithms

for mapping smallholder farms. The crop maps derived in this study are fundamental for crop monitoring, land-use policies and aiding food security planning activities.

Author Contributions: Z.M.-M. conceptualized and developed the original draft of the manuscript. G.J.C. revised the manuscript, supervised and provided financial resources for the project. C.M. was involved in data analysis, reviewed and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This project was funded by the Agricultural Research Council, the National Research Foundation (Grant number: SFH170524232697) and the University of Pretoria.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Sentinel-1 and Sentinel-2 data are freely available from the Copernicus Hub.

Acknowledgments: The researchers would like to express gratitude to the postgraduate students at the Agricultural Research Council: Sabelo Mazibuko, Jillie Masemola and Bonolo Mosuwe for assistance during field data collection. The authors would like to thank the Agricultural Research Council and University of Pretoria for hosting this research. The authors would like to thank the anonymous reviewers for their valuable comments.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Ranum, P.; Peña-Rosas, J.P.; Garcia-Casal, M.N. Global maize production, utilization, and consumption. *Ann. N. Y. Acad. Sci.* **2014**, *1312*, 105–112. [[CrossRef](#)]
- FAO. *Global Information and Early Warning System on Food and Agriculture Crop Prospects and Food Situation*; FAO: Rome, Italy, 2019.
- FAO. *OECD-FAO Agricultural Outlook 2016–2025*; FAO: Rome, Italy, 2016.
- Knox, J.; Hess, T.; Daccache, A.; Wheeler, T. Climate change impacts on crop productivity in Africa and South Asia. *Environ. Res. Lett.* **2012**, *7*, 1–8. [[CrossRef](#)]
- Misra, A.K. Climate change and challenges of water and food security. *Int. J. Sustain. Built Environ.* **2014**, *3*, 153–165. [[CrossRef](#)]
- Giller, K.E.; Rowe, E.C.; de Ridder, N.; van Keulen, H. Resource use dynamics and interactions in the tropics: Scaling up in space and time. *Agric. Syst.* **2006**, *88*, 8–27. [[CrossRef](#)]
- Santpoort, R. The Drivers of Maize Area Expansion in Sub-Saharan Africa. How Policies to Boost Maize Production Overlook the Interests of Smallholder Farmers. *Land* **2020**, *9*, 68. [[CrossRef](#)]
- Homolova, L.; Malenovský, Z.; Clevers, J.G.; García-Santos, G.; Schaepman, M.E. Review of optical-based remote sensing for plant trait mapping. *Ecol. Complex.* **2013**, *15*, 1–16. [[CrossRef](#)]
- Belward, A.S.; Skøien, J.O. Who launched what, when and why; trends in global land-cover observation capacity from civilian earth observation satellites. *ISPRS J. Photogramm. Remote Sens.* **2015**, *103*, 115–128. [[CrossRef](#)]
- Veloso, A.; Mermoz, S.; Bouvet, A.; Le Toan, T.; Planells, M.; Dejoux, J.F.; Ceschia, E. Understanding the temporal behavior of crops using Sentinel-1 and Sentinel-2-like data for agricultural applications. *Remote Sens. Environ.* **2017**, *199*, 415–426. [[CrossRef](#)]
- Drusch, M.; Del Bello, U.; Carlier, S.; Colin, O.; Fernandez, V.; Gascon, F.; Hoersch, B.; Isola, C.; Laberinti, P.; Martimort, P.; et al. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sens. Environ.* **2012**, *120*, 25–36. [[CrossRef](#)]
- Asner, G.P. Cloud cover in Landsat observations of the Brazilian Amazon. *Int. J. Remote Sens.* **2001**, *22*, 3855–3862. [[CrossRef](#)]
- Torbick, N.; Chowdhury, D.; Salas, W.; Qi, J. Monitoring rice agriculture across myanmar using time series Sentinel-1 assisted by Landsat-8 and PALSAR-2. *Remote Sens.* **2017**, *9*, 119. [[CrossRef](#)]
- Cai, Y.; Lin, H.; Zhang, M. Mapping paddy rice by the object-based random forest method using time series Sentinel-1/Sentinel-2 data. *Adv. Space Res.* **2019**, *64*, 2233–2244. [[CrossRef](#)]
- Dobson, M.; Ulaby, F.T.; Pierce, L.E. Land-cover classification and estimation of terrain attributes using synthetic aperture radar. *Remote Sens. Environ.* **1995**, *51*, 199–214. [[CrossRef](#)]
- Van Tricht, K.; Gobin, A.; Gilliams, S.; Piccard, I. Synergistic Use of Radar Sentinel-1 and Optical Sentinel-2 Imagery for Crop Mapping: A Case Study for Belgium. *Remote Sens.* **2018**, *10*, 1642. [[CrossRef](#)]
- Sonobe, R.; Yamaya, Y.; Tani, H.; Wang, X.; Kobayashi, N.; Mochizuki, K. Assessing the suitability of data from Sentinel-1A and 2A for crop classification. *GISci. Remote Sens.* **2017**, *54*, 918–938. [[CrossRef](#)]
- Belgiu, M.; Drăguț, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 24–31. [[CrossRef](#)]
- Cooner, A.J.; Shao, Y.; Campbell, J.B. Detection of urban damage using remote sensing and machine learning algorithms: Revisiting the 2010 Haiti earthquake. *Remote Sens.* **2016**, *8*, 868. [[CrossRef](#)]

20. SDM. *Greater Sekhukhune Cross Border District Municipality INTEGRATED Development Plan: 2019/20*; SDM: Limpopo, South Africa, 2019.
21. Pervez, M.S.; Brown, J.F. Mapping irrigated lands at 250-m scale by merging MODIS data and national agricultural statistics. *Remote Sens.* **2010**, *2*, 2388–2412. [[CrossRef](#)]
22. Lee, J.S.; Jurkevich, L.; Dewaele, P.; Wambacq, P.; Oosterlinck, A. Speckle filtering of synthetic aperture radar images: A review. *Remote Sens. Rev.* **1994**, *8*, 313–340. [[CrossRef](#)]
23. Loew, A.; Mauser, W. Generation of geometrically and radiometrically terrain corrected SAR image products. *Remote Sens. Environ.* **2007**, *106*, 337–349. [[CrossRef](#)]
24. ESA. *Sen2Cor*; ESA: Paris, France, 2018.
25. Jordan, C.F. Derivation of Leaf-Area Index from Quality of Light on the Forest Floor. *Ecology* **1969**, *50*, 663–666. [[CrossRef](#)]
26. Tucker, C.J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150. [[CrossRef](#)]
27. Huete, A.R. A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* **1988**, *25*, 295–309. [[CrossRef](#)]
28. Crippen, R.E. Calculating the vegetation index faster. *Remote Sens. Environ.* **1990**, *34*, 71–73. [[CrossRef](#)]
29. Roujean, J.-L.; Breon, F.-M. Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. *Remote Sens. Environ.* **1995**, *51*, 375–384. [[CrossRef](#)]
30. Chen, J.M. Evaluation of Vegetation Indices and a Modified Simple Ratio for Boreal Applications. *Can. J. Remote Sens.* **1996**, *22*, 229–242. [[CrossRef](#)]
31. Gitelson, A.A.; Kaufman, Y.J.; Merzlyak, M.N. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sens. Environ.* **1996**, *58*, 289–298. [[CrossRef](#)]
32. Rondeaux, G.; Steven, M.; Baret, F. Optimization of soil-adjusted vegetation indices. *Remote Sens. Environ.* **1996**, *55*, 95–107. [[CrossRef](#)]
33. Broge, N.H.; Leblanc, E. Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote Sens. Environ.* **2001**, *76*, 156–172. [[CrossRef](#)]
34. Haboudane, D.; Miller, J.R.; Pattey, E.; Zarco-Tejada, P.J.; Strachan, I.B. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sens. Environ.* **2004**, *90*, 337–352. [[CrossRef](#)]
35. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
36. Rodriguez-Galiano, V.F.; Ghimire, B.; Rogan, J.; Chica-Olmo, M.; Rigol-Sanchez, J.P. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS J. Photogramm. Remote Sens.* **2012**, *67*, 93–104. [[CrossRef](#)]
37. Armitage, D.W.; Ober, H.K. A comparison of supervised learning techniques in the classification of bat echolocation calls. *Ecol. Inform.* **2010**, *5*, 465–473. [[CrossRef](#)]
38. Immitzer, M.; Vuolo, F.; Atzberger, C. First experience with Sentinel-2 data for crop and tree species classifications in central Europe. *Remote Sens.* **2016**, *8*, 166. [[CrossRef](#)]
39. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
40. Mountrakis, G.; Im, J.; Ogole, C. Support vector machines in remote sensing: A review. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 247–259. [[CrossRef](#)]
41. Mirik, M.; Ansley, R.J.; Steddom, K.; Jones, D.C.; Rush, C.M.; Michels, G.J.; Elliott, N.C. Remote distinction of a noxious weed (musk thistle: *Carduus nutans*) using airborne hyperspectral imagery and the support vector machine classifier. *Remote Sens.* **2013**, *5*, 612–630. [[CrossRef](#)]
42. Knorn, J.; Rabe, A.; Radeloff, V.C.; Kuemmerle, T.; Kozak, J.; Hostert, P. Land cover mapping of large areas using chain classification of neighboring Landsat satellite images. *Remote Sens. Environ.* **2009**, *113*, 957–964. [[CrossRef](#)]
43. Huang, C.; Davis, L.S. Townshend JRG. An assessment of support vector machines for land cover classification. *Int. J. Remote Sens.* **2002**, *23*, 725–749. [[CrossRef](#)]
44. Kumar, P.; Gupta, D.K.; Mishra, V.N.; Prasad, R. Comparison of support vector machine, artificial neural network, and spectral angle mapper algorithms for crop classification using LISS IV data. *Int. J. Remote Sens.* **2015**, *36*, 1604–1617. [[CrossRef](#)]
45. Wolpert, D.H. Stacked generalization. *Neural Netw.* **1992**, *5*, 241–259. [[CrossRef](#)]
46. Millard, K.; Richardson, M. On the importance of training data sample selection in random forest image classification: A case study in peatland ecosystem mapping. *Remote Sens.* **2015**, *7*, 8489–8515. [[CrossRef](#)]
47. Zhao, H.; Chen, Z.; Jiang, H.; Jing, W.; Sun, L.; Feng, M. Evaluation of three deep learning models for early crop classification using sentinel-1A imagery time series—A case study in Zhanjiang, China. *Remote Sens.* **2019**, *11*, 2673. [[CrossRef](#)]
48. Olofsson, P.; Foody, G.M.; Stehman, S.V.; Woodcock, C.E. Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. *Remote Sens. Environ.* **2013**, *129*, 122–131. [[CrossRef](#)]
49. Congalton, R.G.; Green, G. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, 2nd ed.; CRC Press: Boca Raton, FL, USA, 2008. [[CrossRef](#)]
50. Efron, B. Estimating the error rate of a prediction rule: Improvement on cross-validation. *J. Am. Stat. Assoc.* **1983**, *78*, 316–331. [[CrossRef](#)]
51. Kuhn, M.; Wing, J.; Weston, S.; Williams, A.; Keefer, C.; Engelhardt, A.; Cooper, T. Caret: Classification and Regression Training, R Package Version 6.0-76. 2017. Available online: <http://cran.r-project.org/package=caret> (accessed on 2 June 2012).

52. McNemar, Q. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika* **1947**, *12*, 153–157. [[CrossRef](#)] [[PubMed](#)]
53. Abubakar, G.A.; Wang, K.; Shahtahamssebi, A.; Xue, X.; Belete, M.; Gudo, A.J.A.; Mohamed Shuka, K.A.; Gan, M. Mapping Maize Fields by Using Multi-Temporal Sentinel-1A and Sentinel-2A Images in Makarfi, Northern Nigeria, Africa. *Sustainability* **2020**, *12*, 2539. [[CrossRef](#)]
54. Useya, J.; Chen, S. Exploring the Potential of Mapping Cropping Patterns on Smallholder Scale Croplands Using Sentinel-1 SAR Data. *Chin. Geogr. Sci.* **2019**, *29*, 626–639. [[CrossRef](#)]
55. Duguay, Y.; Bernier, M.; Lévesque, E.; Tremblay, B. Potential of C and X band SAR for shrub growth monitoring in sub-arctic environments. *Remote Sens.* **2015**, *7*, 9410–9430. [[CrossRef](#)]
56. Khosravi, I.; Safari, A.; Homayouni, S. Separability analysis of multifrequency SAR polarimetric features for land cover classification. *Remote Sens. Lett.* **2017**, *8*, 1152–1161. [[CrossRef](#)]
57. Inoue, Y.; Kurosu, T.; Maeno, H.; Uratsuka, S.; Kozu, T.; Dabrowska-Zielinska, K.; Qi, J. Season-long daily measurements of multifrequency (Ka, Ku, X, C, and L) and full-polarization backscatter signatures over paddy rice field and their relationship with biological variables. *Remote Sens. Environ.* **2002**, *81*, 194–204. [[CrossRef](#)]
58. Ouzemou, J.E.; El Harti, A.; Lhissou, R.; El Moujahid, A.; Bouch, N.; El Ouazzani, R.; Bachaoui, E.M.; El Ghmari, A. Crop type mapping from pansharpened Landsat 8 NDVI data: A case of a highly fragmented and intensive agricultural system. *RSASE* **2018**, *11*, 94–103. [[CrossRef](#)]
59. Sonobe, R.; Tani, H.; Wang, X.; Kobayashi, N.; Shimamura, H. Parameter tuning in the support vector machine and random forest and their performances in cross-and same-year crop classification using TerraSAR-X. *Int. J. Remote Sens.* **2014**, *35*, 7898–7909. [[CrossRef](#)]
60. Foody, G.M.; Mathur, A. The use of small training sets containing mixed pixels for accurate hard image classification: Training on mixed spectral responses for classification by a SVM. *Remote Sens. Environ.* **2006**, *103*, 179–189. [[CrossRef](#)]
61. Thanh Noi, P.; Kappas, M. Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using Sentinel-2 imagery. *Sensors* **2018**, *18*, 18. [[CrossRef](#)]
62. Forkuor, G.; Conrad, C.; Thiel, M.; Ullmann, T.; Zoungrana, E. Integration of optical and Synthetic Aperture Radar imagery for improving crop mapping in Northwestern Benin, West Africa. *Remote Sens.* **2014**, *6*, 6472–6499. [[CrossRef](#)]
63. Deschamps, B.; McNairn, H.; Shang, J.; Jiao, X. Towards operational radar-only crop type classification: Comparison of a traditional decision tree with a random forest classifier. *Can. J. Remote Sens.* **2012**, *38*, 60–68. [[CrossRef](#)]
64. Inglada, J.; Vincent, A.; Arias, M.; Marais-Sicre, C. Improved early crop type identification by joint use of high temporal resolution SAR and optical image time series. *Remote Sens.* **2016**, *8*, 362. [[CrossRef](#)]
65. Arias, M.; Campo-Bescós, M.Á.; Álvarez-Mozos, J. Crop Classification Based on Temporal Signatures of Sentinel-1 Observations over Navarre Province, Spain. *Remote Sens.* **2020**, *12*, 278. [[CrossRef](#)]
66. McNairn, H.; Champagne, C.; Shang, J.; Holmstrom, D.; Reichert, G. Integration of optical and Synthetic Aperture Radar (SAR) imagery for delivering operational annual crop inventories. *ISPRS J. Photogramm. Remote Sens.* **2009**, *64*, 434–449. [[CrossRef](#)]
67. Sonobe, R.; Yamaya, Y.; Tani, H.; Wang, X.; Kobayashi, N.; Mochizuki, K.I. Crop classification from Sentinel-2-derived vegetation indices using ensemble learning. *J. Appl. Remote Sens.* **2018**, *12*, 026019. [[CrossRef](#)]
68. Yi, Z.; Jia, L.; Chen, Q. Crop Classification Using Multi-Temporal Sentinel-2 Data in the Shiyang River Basin of China. *Remote Sens.* **2020**, *12*, 4052. [[CrossRef](#)]
69. Zhang, H.; Kang, J.; Xu, X.; Zhang, L. Accessing the temporal and spectral features in crop type mapping using multi-temporal Sentinel-2 imagery: A case study of Yí'an County, Heilongjiang province, China. *Comput. Electron. Agric.* **2020**, *176*, 105618. [[CrossRef](#)]
70. SDG. *Sustainable Development Goals*; United Nations: New York, NY, USA, 2019.
71. Foody, G.M. Sample size determination for image classification accuracy assessment and comparison. *Int. J. Remote Sens.* **2009**, *30*, 5273–5291. [[CrossRef](#)]
72. Forkuor, G.; Dimobe, K.; Serme, I.; Tondoh, J.E. Landsat-8 vs. Sentinel-2: Examining the added value of sentinel-2's red-edge bands to land-use and land-cover mapping in Burkina Faso. *GISci. Remote Sens.* **2018**, *55*, 331–354. [[CrossRef](#)]
73. Kim, H.O.; Yeom, J.M. Effect of red-edge and texture features for object-based paddy rice crop classification using RapidEye multi-spectral satellite image data. *Int. J. Remote Sens.* **2014**, *35*, 7046–7068. [[CrossRef](#)]