Exploring machine learning algorithms for mapping crop types in a heterogeneous agriculture landscape using Sentinel-2 data. A case study of Free State Province, South Africa.

[2]Theresa Taona Mazarire, [1,2]Phathutshedzo Eugene Ratshiedana, [1]Adolph Nyamugama, [2]Elhadi Adam, [1]George Chirima.

[1]Geo-information Division, Institute for soil water and climate, Agricultural Research Council (ISCW-ARC) Postal address: Private Bag x 79, Pretoria, 001 South Africa.
NyamugamaA@arc.agric.za
[2]Department of Geography, Archaeology and Environmental, University of Witwatersrand. Private Bag x3, Wits 2050, Johannesburg, South Africa.

DOI: http://dx.doi.org/10.4314/sajg.v9i2.22

Abstract

Accurate and detailed studies in crop mapping are crucial in precision agriculture, yield estimations, and crop monitoring. This study focused on exploring the utility of Sentinel-2 data in mapping of crop types and testing the two machine learning algorithms which are Random Forest and Support Vector Machine performance in classifying crop types in a heterogeneous agriculture landscape in Free state province, South Africa. Nine crop types were successfully classified. The utility and contribution of different bands for classification were evaluated using RF mean decrease GINI for variable importance. Validation of results was done using a confusion matrix which produced overall accuracy, errors and prediction measures. The best performance was attained by SVM with an overall accuracy of 95% and a kappa value of 94%. RF also performed fairly well with 85% of overall accuracy and kappa value of 83%. It was concluded that Sentinel-2 data performs better using the SVM classifier compared to RF classifier.

***Keywords**: Sentinel-2 Multispectral Instrument (MSI), (SVM) Support vector machine, (RF) Random Forest; machine learning algorithms, Remote sensing, Image processing, heterogeneous landscape.*

1. Introduction

Globally food security remains a major challenge especially in the developing countries (Shafiee and Cai, 2016). This is a consequence of population expansion as well as the impacts of climate change in food production (Altieri and Nicholls, 2017). Although the rest of the world has made substantial progress towards alleviating food insecurity, Africa, especially Sub-Saharan Africa remains behind (Blein *et al*., 2013). Agriculture is the backbone of South Africa's economy and it has grown by 22% in the last 5 years contributing 2.2% to the country's GDP (Boshoff and Fourie, 2020). In spite of this, the agricultural sector in the country has not been able to meet the demand for the main food items consumed domestically since 2000 (Greyling, 2012; Cai *et al*., 2017). To meet the unending future food demands, there is a need for effective agriculture management and planning to understand crop dynamics and distribution. This involves using accurate, reliable and inclusive agriculture intelligence technology to better manage the agriculture landscape and map crops (Löw, 2013).

Crop type mapping has been traditionally carried out using a field-based survey. However, census and the ground-based survey are not feasible in a large scale agriculture landscape as it is laboriously expensive and prone to errors and time consuming hence the use of timely, less expensive and faster methods of crop mapping has become a necessity (Ouzemou *et al*., 2018, Defournya, 2019). Remote sensing (RS) technology has emerged as an effective and important source of land use information allowing enhancing broad-scale agronomic management and within field monitoring such as precision agriculture. This technological advancement gives crop information which enables effective Spatio-temporal monitoring of crops as well as providing valuable data used to identify crop types and their corresponding location, the nutrient status of each crop, expected crop yield and assessment of actual crop distribution and extent of damage thus saving time and resources (Asgarian *et al*., 2016).

In South Africa, the use of remote sensing technology in agriculture is particularly challenging as most of the communal landscape is characterized by fragmented, small parcel sized fields and different crop types within a pixel as well as highly heterogeneous crop cover (Cai *et al*., 2017). This presents a limitation because different crop types have high variability in phenological stages within fields such as early sprouting, establishment and maturation which might lead to the same spectral signature at some point in their development (Asgarian *et al*., 2016). In the absence of distinct crop border, crop type identification becomes a challenge as there is need to employ expensive high spectral resolution to identify features which distinguish one crop from another (Veloso *et al*., 2017).

The availability of the satellite-based imaging technology provides multi-spectral data which is collected on regular revisiting intervals and free from weather disturbances which have made the mapping of crop types possible at all seasons (Li *et al*., 2014). High spatial resolution sensors such as SPOT and ASTER have been used to accurately differentiate between agriculture crops (Conrad *et al*., 2010; Inglada *et al*., 2015; Gilbertson and Van Niekerk, 2017). However, high-resolution data is costly and has low temporal revisits. The introduction of free and easily accessible Sentinel-2 multi-spectral instrument (MSI) gives new possibilities for classification of crop types.

The improved characteristics of Sentinel-2 meet the requirements for crop mapping as it has shorter revisit periods to cater for the temporal dynamics of crop growth together with medium spatial resolution over wide spectral channels enabling the discrimination of acute vegetation spectral signatures (Malenovský *et al*., 2012, Defournya, 2019). Additionally, it possesses high radiometric and spectral resolution that allows acute vegetation properties to be identified. MSI data with a medium spatial resolution has been previously used for crop mapping (Conrad *et al*., 2010, Li *et al*., 2014, Ouzemou *et al*., 2018, Defournya,2019).

The development of machine learning algorithms offers an interesting prospect to further investigate how the classifiers perform in a heterogeneous agriculture landscape. Besides selecting the appropriate RS data for the area, the choice of classification method is also equally important for a successful crop type mapping (Sothe *et al*., 2017). The most commonly used machine learning algorithms in heterogeneous land cover classification include Decision trees, Artificial Neural Network (ANN), Random Forest (RF), and Support Vector Machine (SVM) amongst others. Nonparametric classifiers are considered suitable and superior because of their ability to ignore the assumption of a normal distribution for the dataset and also statistical parameters are not required when separating image classes (Adam *et al*., 2014; Inglada *et al*., 2015; Sothe *et al*., 2017; Ouzemou *et al*., 2018). As such, modern algorithms have been recommended (RF and SVM) as they overcome the shortcomings of the traditional algorithms because they can synthesize classification functions using either discrete or continuous datasets (Sothe *et al*., 2017). They are also insensitive to noise which makes them not to be constrained by parametric distribution assumptions (Löw *et al*., 2013). Particularly, SVM classifier with the use of Gaussian Kernal density function has emerged superior in most studies (Nitze *et al*., 2012; Inglada *et al*., 2015, Kumar *et al*., 2017), although other studies report that RF performs marginally better than SVM in heterogeneous landscape classification (Adam *et al*., 2014). However, there is a gap as to which classifier is suitable to distinctively classifying crop types in particular.

This background motivated the present study to investigate the performance of Sentinel-2 MSI data as well as evaluating the performance of RF and SVM for mapping crop types in a highly heterogeneous agriculture landscape of Free State province in South Africa. The province contributes a significant input to the country's food security, hence need to promote and invest in agriculture technology.

2. Methods and materials

2.1. Study Area

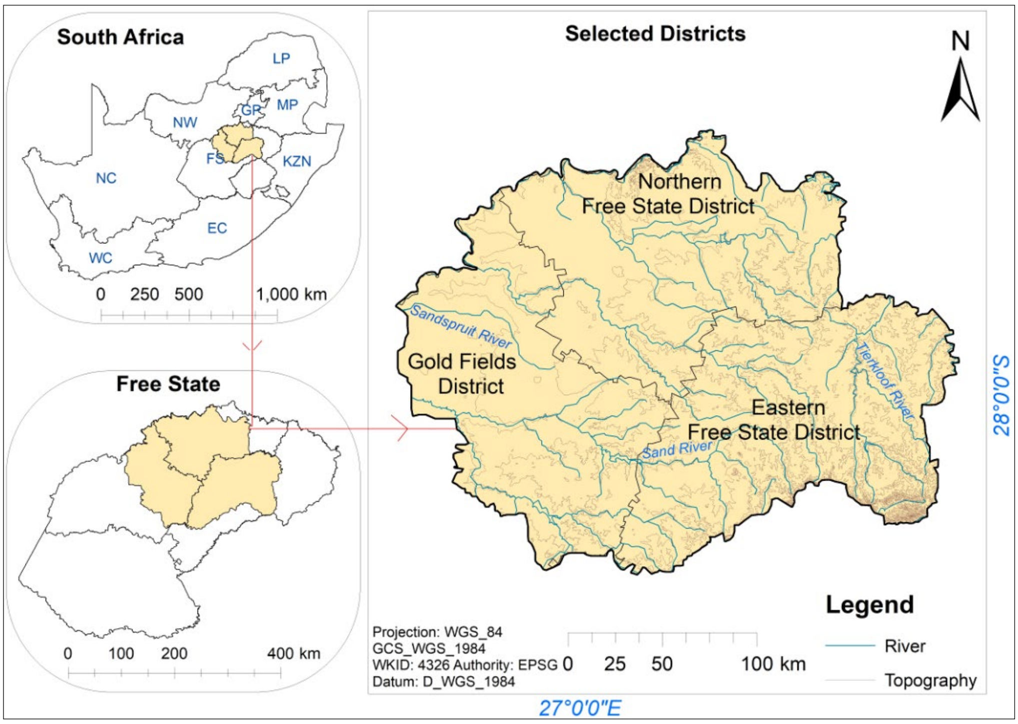

Figure 1: Location of study area indicated in topaz sand colour in South Africa and within Free State province districts.

The study area is located in the middle to the upper part of Free State, South Africa and forms part of the former QwaQwa homeland. Farmers in this area apply the principle of crop rotation and irrigation scheme to enhance the agriculture intensity (Butler *et al*., 1978).

2.2. Data acquisition and pre-processing

This section presents the methods, procedures and data employed during the study until the generation of results. The work was divided into four stages, namely; field data collection, use of spot 6/7 for visualization with sentinal-2 imagery, pre-processing of imageries and classification of R-studio using two classifiers, see figure 2. Section 2.2.1 to 2.3.3; explain in details the methodologies used in this study.

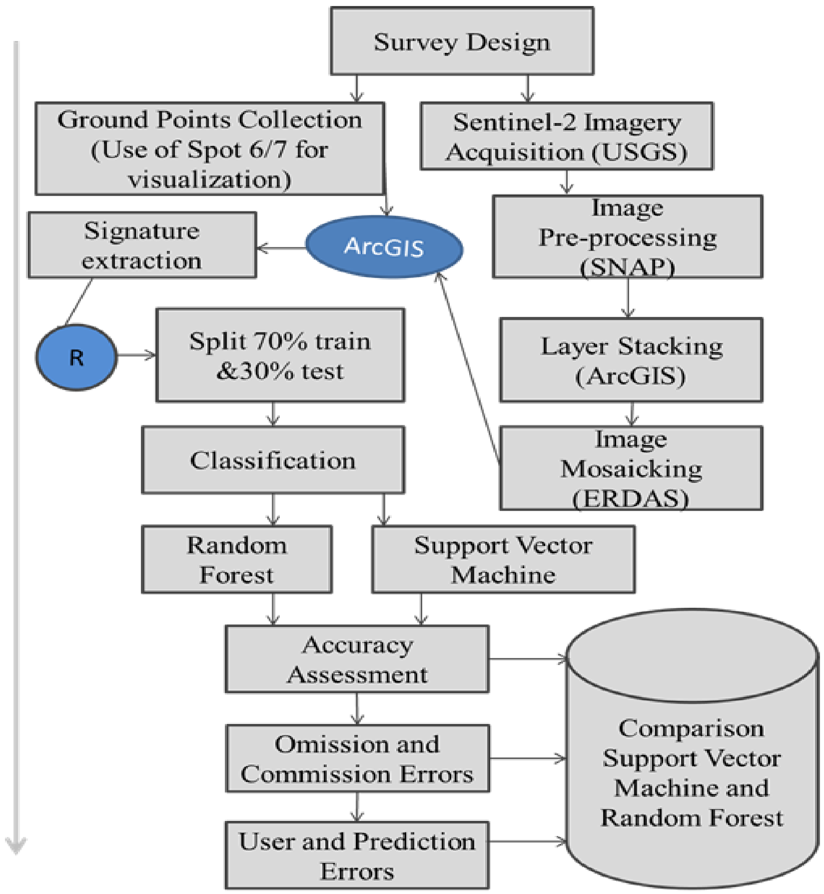

Figure 2: Steps followed from the study design throughout data collection and processing.

2.2.1. Field data collection

The ground control points were randomly collected through a field survey conducted in February 2019. A Global Positioning System receiver (GPS) was used to record the longitude and latitude coordinates of the different crop types and land use classes in the study area. More ground points were manually digitized using visual interpretation of SPOT 6/7 on Google earth and Sentinel-2 MSI imagery. The ground control points were split into 70% training and 30% validation for subsequent classification.

2.2.2. Satellite data acquisition and pre-processing

To provide blanket coverage of the study area, nine Sentinel-2 Multi-Spectral sensor imageries were acquired for February 2019 with a cloud cover of less than 10%. The images were obtained from the USGS' Earth Explorer portal (https://earthexplorer.ugs.gov/). The acquisition dates correspond to the sprouting and or harvesting time of the crops in order to make discrimination between the crops easier.

2.2.3. Image pre-processing

The imageries were atmospherically corrected using the Level 2A prototype processor (Sen2Cor) which is a plugin in the SNAP software package. The generated output images were resampled and generated with an equal spatial resolution (10m) for all bands. The corrected bands were stacked chronologically into one image for further analysis.

All images were individually cleaned to remove speckle and noise using edge trimming, haze reduction and cloud masking. To achieve radiometric equivalence for all nine scenes, histogram matching was done using the best scene as a reference scene. The procedure was performed band to band using a mathematically determined lookup table to convert the histogram of one scene to be like the histogram of the reference scene (ERDAS, 1997). Spectral reflectance values of the ground points were extracted from the image and used to model the classifiers.

2.3. Image classification

Support Vector Machine and Random Forest algorithms were used to classify the area. These two classifiers were selected due to their high performance in classifying crop types and heterogeneous landscapes (Nitze *et al*., 2012).

2.3.1. Support Vector Machine classification

Support Vector Machine is a supervised nonparametric technique which is trained to find an optimal classification hyperplane by grouping classes based on the statistical learning theory thus there is no assumption on the underlying data distribution (Mountrakis *et al*., 2011). The hyperplane in SVM was developed using the training data (70%) and validated using the independent testing data set (30%). Radial basis function was selected because of its popularity in remote sensing studies as it outperforms the other kernels (Pal and Mather, 2005, Inglada *et al*., 2015, Kumar *et al*., 2017). The equation for the Radial Basis Function is as follows (Kumar *et al*., 2017):

$$K{:}\,(x, xj) = (-y \parallel (xi, xy) \parallel 2), y > 0 \qquad [1]$$

Where y is the gamma which controls the width of the Gaussian kernel function. The accuracy of the SVM algorithm is determined by algorithm parameterization which requires the regularization of parameter C-cost and definition of the kernel parameter Y- gamma (Oommen *et al*., 2008). 10-fold cross-validation was utilized which overcomes the overfitting problems (Huang *et al*., 2002). Pairs of Y and C were tested hence parameters with the best performance were automatically used for training the SVM model. The optimized SVM parameter values were obtained from gamma value 0.33 and cost value 100 which was then used to classify the images. The SVM model was run in R statistical software using caret and e1071 package (Hornik *et al*., 2006).

2.3.2. Random Forest classification

Random Forests are an ensemble of tree predictors that depend on the value of independent random vector sampled for all the trees in the forest (Breiman, 2001). The algorithm is combined with many ensemble regression and classification trees to build binary classification trees using several bootstrap samples with other trees drawn from the original dataset. It is superior to many tree-based algorithms since it is not sensitive to noise, highly accurate, robust to outliers as well as the calculation of different internal diagnostic indicators like out of bag (OOB) and variable importance thus not subject to over-fitting (Breiman, 2001). RF constructs many decision trees using the training dataset (70%) which is a random, bootstrapped subset hence keeping the remaining dataset (30%) for the internal error testing (Nitze *et al*., 2015).

Two parameters were defined to initiate the algorithm, the number of trees used in the forest (ntree) and several random variables used in each tree (mtry) (Breiman, 2001). For improved accuracy, the parameters (mtry and ntree) that provide sufficiently low correlation with adequate predictive power were optimized using a grid search approach. The optimum results were obtained by setting the number of variables (m) equal to the square root of the number of overall variables (M) as suggested by Breiman (2002). The algorithm makes use of the Classification and Regression Trees (CART) to generate trees (Breiman, 2001), hence the tree node is split according to a criterion. The GINI index criterion which measures the impurity of a given element concerning the rest of the classes (Rodriguez-Galiano *et al*., 2012), was used to perform the split. GINI Index can be written as the equation below (Ok *et al*., 2012):

$$\sum\sum_{j\neq i}\left(\frac{\mathrm{f(ci,T)}}{|\mathrm{T}|}\right)\left(\frac{\mathrm{f(ci,T)}}{|\mathrm{T}|}\right) \qquad [2]$$

Where T is the 70% training set, Ci is the class that a randomly selected pixel belongs to f(Ci, T) |T| is the probability that the selected case belongs to class Ci. The relative importance of different variables was calculated using the mean decrease during the Out-of-bag (OOB) error calculation.

2.3.3. Accuracy assessment

The accuracy assessment was done on the classified image in order to evaluate the ability of the sensor in discriminating the crop types using the 30% validation data. Confusion matrices were constructed to compute the overall accuracy, kappa coefficient, omission error, commission error, user accuracy and producer's accuracy for both classifiers.

3. Results

3.1. Optimization of the algorithms

The classification algorithms were optimized to select the best parameters to. SVM used 10-fold cross-validation which yielded a gamma (Y) of 1, with the lowest error produced from the combination of gamma (γ) value of 0.1 and cost (C) value of 100 (Figure 3 a). The RF input parameters of mtry value of 4 combined with ntree value of 3500 produced the lowest OOB error rate of 21.5% and mtry and ntree values of 8 and 4500 respectively produced the highest OOB error rate of 25 %.(Figure 3 b).

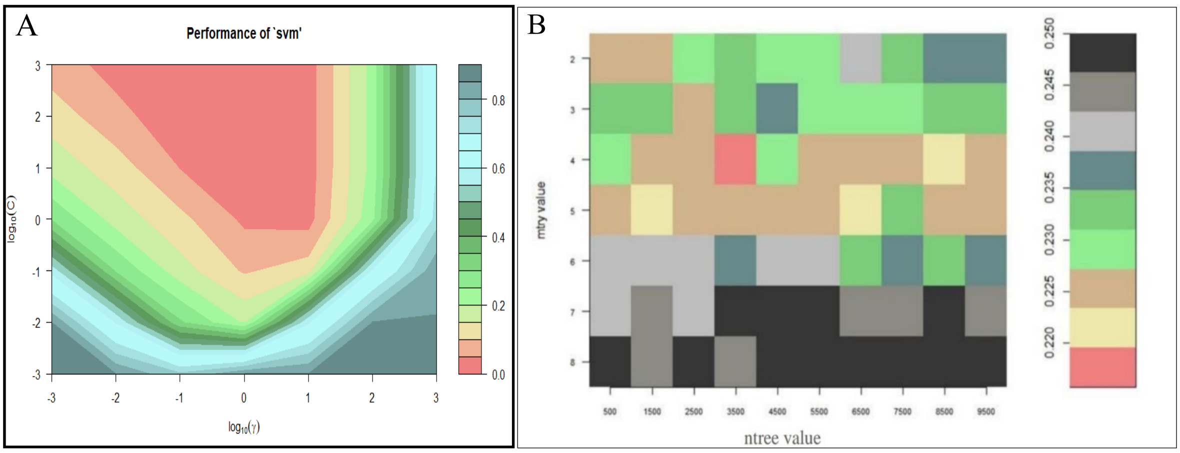

Figure 3: (A) SVM parameter optimization (B) RF optimization

3.2. Feature importance determined by random forests

While the primary goal of the RF analysis was to identify feature subsets for classification and to compare it with the SVM classification, interesting conclusions can also be drawn from the RF results. RF was used to provide a variable importance measurement to indicate every band's contribution to the agriculture landscape mapping. The most important bands were allocated in blue, SWIR1, SWIR2 and all the vegetation red edge bands. This result is likely due to the dominant crop cover type in the region which is maize.

The two classification methods used in this study generally showed a reasonable accurate visual depiction of the heterogeneous agricultural land use. Nine crop types were accurately classified (Figure 4). Similar to the study of Myburgh and Van Niekerk (2013), confusion matrix was computed for both classifiers to obtain Overall Accuracy (OA), Kappa Coefficient and errors for accuracy assessment. RF performance was observed to be low with 85% overall accuracy compared to SVM which performed higher with 95% overall accuracy. To complement the accuracy of the two, RF produced a kappa of 0.83 while SVM produced a kappa coefficient of 0.94. Both classifiers were computed at 95 % confidence intervals.

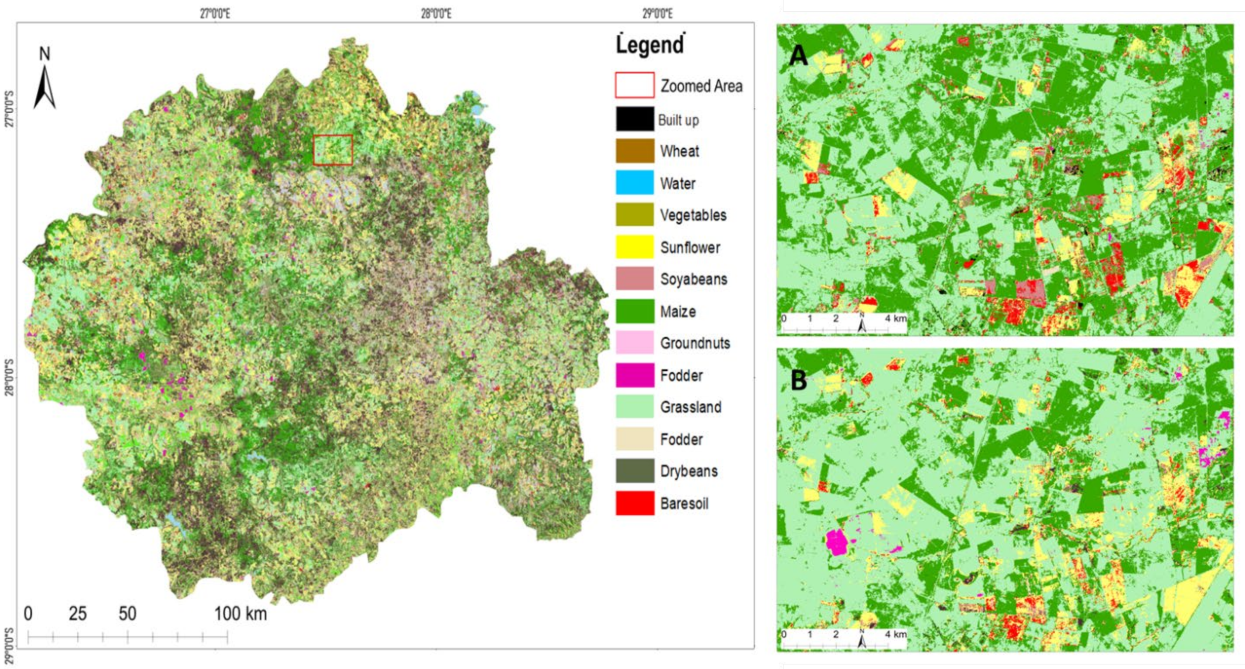

Figure 4: Classification maps showing crop types in Free State upper region- February 2019 using (a) SVM algorithm (b) RF algorithm.

Also, commission error (CE) and omission error (OE) percentages were computed for all classes (Figure 5.a-c). The omission error indicates the areas which were misclassified while the commission error percentage indicates the misclassification of values which have been predicted in another class which they do not belong. The results show that RF has more errors compared to SVM making SVM a better classifier with fewer errors. This is evident with bare soil, dry beans, grassland, sorghum and vegetables have a high percentage of omission error. The commission errors are high in RF classifier compared to SVM. However, both classifiers showed no error in classifying wheat (Table 1 and Table 2).

To further assess the accuracy, analysis based on the user and the predictors were also computed (Table 1 and Table 2). The assessment based upon the user's accuracy (UA) and predictor's accuracy (PA) revealed that SVM based on prediction accuracy (PA) outplays the RF classifier except in classifying lucerne where it is lower and equal on maize classification, on the other hand, the similarities occur also in user accuracy were SVM performs better than RF on average (Table 1 and Table 2).

Table 1: Random Forest Statistical Results.

Land Cover	Omission Errors RF%	Commission Errors RF%	Prediction Accuracy RF	User Accuracy RF
Bare soil	22.12	12.04	56.89	87.96
Dry beans	3.39	23.53	88.64	76.47
Fodder	31.25	45	23.91	55
Grassland	12.63	6.67	82.35	93.33
Groundnuts	2.52	5.30	92.86	94.70
Lucerne	16.67	2.53	89.53	97.47
Maize	0	4.42	100	95.58
Sorghum	2.10	8.20	95.73	91.80
Sunflower	0	35.53	98.99	64.47
Vegetables	3.03	28.38	96.47	71.62
Water	0	0.87	99.13	99.13
Wheat	0	0	91.89	100
OA %	**Kappa**			
84.57	0.83			

Table 2: Support Vector Machine Statistical Results.

Land Cover	Omission Error SVM%	Commission Errors SVM%	Prediction Accuracy SVM	User Accuracy SVM
Bare soil	43.11	12.90	77.89	87.10
Dry beans	11.36	7.57	96.61	92.43
Fodder	10.87	47.62	68.75	52.38
Grassland	17.65	2.35	87.37	97.65
Groundnuts	7.14	1.90	97.48	98.10
Lucerne	16.28	5.41	83.33	94.59
Maize	0	2.22	100	97.78
Sorghum	2.56	4.11	97.90	95.89
Sunflower	0	1.11	100	98.89
Vegetables	3.53	5.88	96.97	94.12
Water	0	0	100	100
Wheat	0	0	100	100
OA%	**Kappa**			
94.62	0.94			

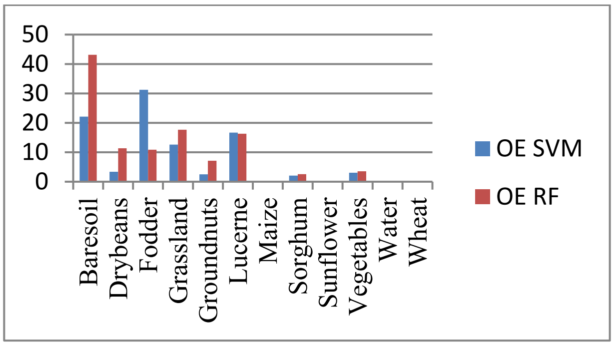

Figure 5. a: Omission Errors (SVM&RF)

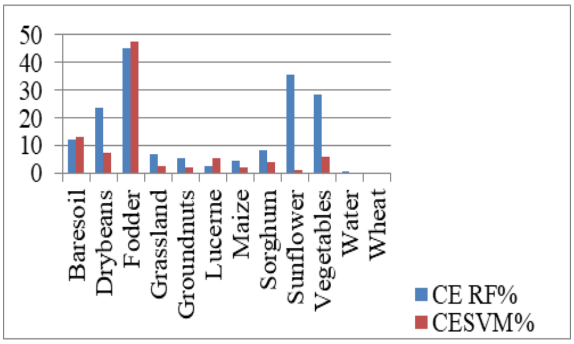

Figure 5. b: Commission Errors (SVM&RF)

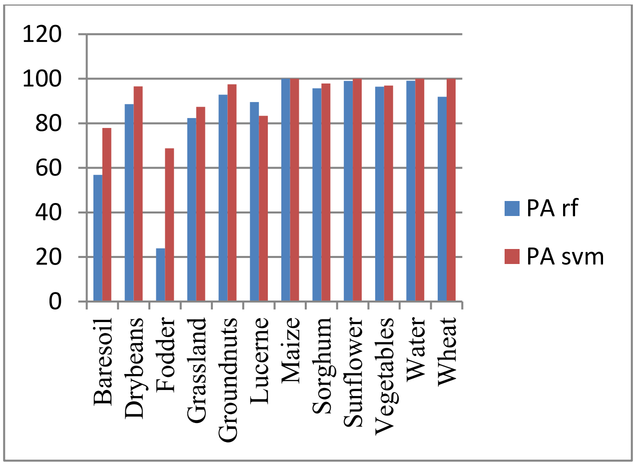

Figure 5.c: Prediction Accuracy (SVM&RF)

Table 3: Area classified by SVM and RF classifiers based on Sentinel 2 MSI data.

Land Cover Type	SVM Area (ha)	RF Area (ha)	% Difference
Non-Crop	175094	535053	43.17826
Maize	1670418	1410916	31.12811
Wheat	601670	481020	14.47236
Vegetables	324506	301965	2.703867
Sunflower	597639	621721	2.888715
Soybeans	157295	162312	0.601806
Sorghum	117347	134464	2.05324
Groundnuts	180748	191402	1.277982
Lucerne	195310	183479	1.419167
Dry bean	28201	25896	0.276492
Total	4048228	4048228	100

As shown in Table 3, the area was dominated by maize (1670418 ha) in SVM and 259502 ha in RF, dry bean classified the lowest by both algorithms, 28201 ha for SVM and 25896 ha for RF. There is less area unclassified by the SVM algorithm compared to RF. This is an indication that SVM classifier classified best compared to the RF classifier. For this statistical analysis, other classes such as water, bare rock, bare soil and other artificial land cover types were removed since the main objective of the study is to classify crop types.

4. Discussions

The present study sought to utilize the recently launched Sentinel-2 missions to evaluate the machine learning algorithms RF and SVM performance in discriminating heterogeneous agriculture landscape. Past researchers have used this available data and augmented their experiments by using vegetation indices and phenological information from long term field survey (Asgarian *et al*., 2016) and with the use of pixel-based classification (Immitzer *et al*., 2016, Vuolo *et al*., 2018), and object-based image analysis (Schultz *et al*., 2015). This study aimed to demonstrate that crop types can also be accurately mapped without using vegetation indices like the NDVI and phenology concept. The results of this study show that in places where ground data is unavailable, low spatial resolution multispectral data can be used to digitize the location of the crop types. This study has proved that supervised models can be trained to classify heterogeneous areas, which become less costly compared to other manual mapping techniques.

The results generated from the methodology includes digital maps, tables and graphs derived from Sentinel-2 MSI imagery classification, where; variable importance and accuracy assessment were evaluated to determine the best method and bands to use when classifying crop type in a heterogeneous landscape. From the classified maps (Figure 4), it is evident that the area has been classified into 12 classes for February 2019 and these classes were detected and distinguished accurately. Concerning the multispectral data used for this study, Sentinel-2 MSI obtained comparatively high classification accuracies showing the considerable capability of the moderate spatial resolution data in accurately identifying and distinguishing the different crop types and land use classes with areas in hectares (Table 1 and Table 2).

The study tested two familiar machine learning, RF and SVM in crop identification. The classification results demonstrated that both RF and SVM algorithms are valuable in mapping and understanding complex agriculture landscapes as both classifiers produced equally good overall accuracies (Table 1 and Table 2) as noted by Nitze *et al*., 2012; Ouzemou *et al*., 2018). The importance of each of the bands in each sensor on the classification output was successfully determined in the study by utilizing RF's variable importance. Congalton and Green (2008), postulates that this is an important contribution for resolving classification errors which are often associated with utilizing multi-spectral imagery. The study demonstrated the value of each band in enhancing the accuracy of crop type classification classes. It also confirmed the high contribution of the red edge and shortwave infrared (SWIR) bands in crop type mapping as postulated by (Immitzer *et al*. 2016) in vegetation and crop mapping.

Regarding the performance of the classifier, it was observed that SVM produced slightly higher classification accuracy than RF when applied to Sentinel-2 data. The present study found that nonlinear kernel function is efficient in SVM classification because it can solve inseparability problems that may be found in land use classes in a heterogeneous landscape. This result agrees with the findings of Löw *et al*. (2013), who examined the influence of feature space size on SVM and RF performance for field-based crop classification using multispectral Rapid Eye data. The study also agrees with the work of Gilbertson and Van Niekerk (2017), wherein they studied the value of dimensionality reduction for crop differentiation with multi-temporal imagery and machine learning and concluded that SVM algorithm can be used to the full set of features generated.

5. Conclusions

The present study assessed the capability of the relatively recent and free sensors, with advanced machine learning classifiers - RF and SVM - to identify and discriminate crop types from the co-existing land-use types in a highly heterogeneous agriculture environment. The outcome of this research provides local confirmation of the performance of Sentinel-2 for classifying crop types for a complex regional landscape. Sentinel-2 data achieved superior results showing that the vegetation red-edge bands are noteworthy thus agreeing with other researches that also indicated the importance of these spectral bands in vegetation classification. SVM proved to be flexible as the same parameters were used in all experiments. The utilization of RF algorithm allowed for rigorous evaluation of the band importance for classification of crop types and it was easy to manipulate because the optimized parameters were minimal. In conclusion, this study showed that the adopted approaches to using machine learning algorithms for crop type mapping were promising. The results of this study are relevant not only for crop monitoring and planning for food security management but also for subsidizing actions for precision agriculture for the communal farmers in South Africa without forgetting that the study allows the mapping of small scale farms also to know the farm boundaries and their yields allowing the right distribution of inputs to smallholders.

6. Recommendations

The interpretation of results in this study can only be regarded as preliminary, therefore; further research is needed to widen the use of multi-spectral imagery on mapping crop types in a heterogeneous landscape. There is a need for further research to develop a technique capable of accurately analysing and discriminating the different crop types found in the small agriculture fields in the African landscape. Furthermore, the use of high spatial resolutions data results in misclassification of the crop types that have the same high spectral characteristics, especially within small fields. Therefore, further investigation is needed on the use of medium resolution data by employing other classification approaches like object-based classification technique. Additionally, the study recommends the use of diffusion data of multispectral data and radar data such as Sentinel-1 for crop type mapping.

7. Acknowledgements

This research was supported by the Agricultural Research Council (ARC) Climate Change Collaboration Centre and SEN2-Agri project by ESA. We thank our colleagues from the University of Witwatersrand- Department of Geography towards the production of this paper.

8. References

Adam, E., Mutanga, O., Odindi, J. & Abdel-Rahman, E. M. 2014. Land-use/cover classification in a heterogeneous coastal landscape using RapidEye imagery: evaluating the performance of random forest and support vector machines classifiers. *International Journal of Remote Sensing*, 35, 3440-3458.

Altieri, M.A. and Nicholls, C.I., 2017. The adaptation and mitigation potential of traditional agriculture in a changing climate. *Climatic Change*, 140(1), pp.33-45.

Asgarian, A., Soffianian, A. and Pourmanafi, S. 2016. Crop type mapping in a highly fragmented and heterogeneous agricultural landscape: A case of central Iran using multi-temporal Landsat 8 imagery. *Computers and Electronics in Agriculture*, 127, 531-540.

Blein, R., Bwalya, M., Chimatiro, S., Faivre-Dupaigre, B., Kisira, S., Leturque, H. & WamboYamdjeu, A. 2013. African agriculture, transformation, and outlook. *Johannesburg: New Partnership for African Development*.

Boshoff, W.H. and Fourie, J., 2020. The South African economy in the twentieth century. In *Business Cycles and Structural Change in South Africa* (pp. 49-70). Springer, Cham.

Breiman, L. 2001. Random forests. *Machine learning*, 45, 5-32.

Breiman, L. 2002. Manual on setting up, using, and understanding random forests v3. 1. 2002. *URL: http://oz. berkeley. edu/users/breiman/Using_random_forests_V3, 1.*

Butler, J., Rotberg, R.I. and Adams, J., 1978. The black homelands of South Africa: The political and economic development of Bophuthatswana and Kwa-Zulu (Vol. 396). *Univ of California Press. http://ark.cdlib.org/ark:/13030/ft0489n6d5/*

Cai, Y., Guan, K., Peng, J., Wang, S., Seifert, C., Wardlow, B. and Li, Z., 2018. A high-performance and in-season classification system of field-level crop types using time-series Landsat data and a machine learning approach. *Remote sensing of environment*, 210, pp.35-47. https://doi.org/10.1016/j.rse.2018.02.045

Congalton, R. G. & Green, K. 2008. Assessing the Accuracy of Remotely Sensed Data: Principles and Practices, (Mapping Science). *CRC Press: Boca Raton, FL, USA*.

Conrad, C., Fritsch, S., Zeidler, J., Rücker, G. & Dech, S. 2010. Per-field irrigated crop classification in arid Central Asia using SPOT and ASTER data. *Remote Sensing*, 2, 10351056.

Defournya, P.,Bontempsa, S., Bellemans, N ., Carab, C., Dedieu, G.,Guzzonatod,E, Hagollec, O., Ngladac, J., Nicolab, L., Rabauted, T Savinaudd, M ., Udroiub, C., Valeroc, S., Béguée, A,Jean, f , Dejouxc-F., Hartig A, Ezzaharh J,, Kussuli, N., Labbassij, K., Lebourgeoise, V., Miaok,Z., Newbyl, T., Nyamugamal, A., Salhm, N., Shelestovi, A., t Simonneauxc, V., Sibiry, P., , Souleymane T., .Traorep, S., Koetz ,B., (2019) Near real-time agriculture monitoring at national scale at parcel resolution: Performance assessment of the Sen2-Agri automated system in various cropping systems around the world. Remote sensing of Environment 221, 551-568 .

ERDAS (Firm), 1997. ERDAS field guide. *Erdas*.
http://www.geoservis.ftn.uns.ac.rs/downloads/ISP/FieldGuide.pdf

Gilbertson, J.K. and Van Niekerk, A., 2017. Value of dimensionality reduction for crop differentiation with multi-temporal imagery and machine learning. *Computers and Electronics in Agriculture*, *142*, pp.50-58.

Greyling, J. C. 2012. *The role of the agricultural sector in the South African economy.* Stellenbosch: Stellenbosch University.

Hornik, K., Meyer, D. & Karatzoglou, A. 2006. Support vector machines in *R. Journal of statistical software*, 15, 1-28.

Huang, C., Davis, L. & Townshend, J. 2002. An assessment of support vector machines for land cover classification. *International Journal of remote sensing*, 23, 725-749.

Immitzer, M., Vuolo, F. & Atzberger, C. 2016. First experience with Sentinel-2 data for crop and tree species classifications in central Europe. *Remote Sensing*, 8, 166.

Inglada, J., Arias, M., Tardy, B., Hagolle, O., Valero, S., Morin, D., Dedieu, G., Sepulcre, G., Bontemps, S. & Defourny, P. 2015. Assessment of an operational system for crop type map production using high temporal and spatial resolution satellite optical imagery. *Remote Sensing*, 7, 12356-12379.

Kumar, P., Prasad, R., Choudhary, A., Mishra, V. N., Gupta, D. K. & Srivastava, P. K. 2017. A statistical significance of differences in classification accuracy of crop types using different classification algorithms. *Geocarto International*, 32, 206-224.

Li, L., Friedl, M. A., Xin, Q., Gray, J., Pan, Y. & Frolking, S. 2014. Mapping crop cycles in China using MODIS-EVI time series. *Remote Sensing,* 6**,** 2473-2493.

Löw, F., Michel, U., Dech, S. & Conrad, C. 2013. Impact of feature selection on the accuracy and spatial uncertainty of per-field crop classification using support vector machines. *ISPRS Journal of photogrammetry and remote sensing,* 85**,** 102-119.

Malenovský, Z., Rott, H., Cihlar, J., Schaepman, M. E., García-Santos, G., Fernandes, R. & Berger, M. 2012. Sentinels for Science: Potential of Sentinel-1,-2, and-3 missions for scientific observations of ocean, cryosphere, and land. *Remote Sensing of Environment,* 120**,** 91-101.

Mountrakis, G., Im, J. and Ogole, C. 2011. Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing, 66, 247-259.*

Myburgh, G. and Van Niekerk, A., 2013. Effect of feature dimensionality on object-based land cover classification: A comparison of three classifiers. *South African Journal of Geomatics*, *2*(1), pp.13-27.

Nitze, I., Barrett, B. & Cawkwell, F. 2015. Temporal optimization of image acquisition for land cover classification with Random Forest and MODIS time-series. *International Journal of Applied Earth Observation and Geoinformation*, 34, 136-146.

Ok, A. O., Akar, O. & Gungor, O. 2012. Evaluation of random forest method for agricultural crop classification. *European Journal of Remote Sensing*, 45, 421-432.

Oommen, T., Misra, D., Twarakavi, N. K., Prakash, A., Sahoo, B. & Bandopadhyay, S. 2008. An objective analysis of support vector machine-based classification for remote sensing. *Mathematical Geosciences*, 40, 409-424. https://rdcu.be/bYSWO

Ouzemou, J.-E., El Harti, A., Lhissou, R., El Moujahid, A., Bouch, N., El Ouazzani, R., Bachaoui, E. M. & El Ghmari, A. 2018. Crop type mapping from pansharpened Landsat 8 NDVI data: A case of a highly fragmented and intensive agricultural system. *Remote Sensing Applications: Society and Environment*, 11, 94-103.

Pal, M. & Mather, P. 2005. Support vector machines for classification in remote sensing. *International Journal of Remote Sensing,* 26**,** 1007-1011.

Rodriguez-Galiano, V. F., Ghimire, B., Rogan, J., Chica-Olmo, M. & Rigol-Sanchez, J. P. 2012. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 67, 93-104.

Schultz, B., Immitzer, M., Formaggio, A., Sanches, I., Luiz, A. & Atzberger, C. 2015. Self-guided segmentation and classification of multi-temporal Landsat 8 images for crop type mapping in southeastern Brazil. *Remote Sensing*, 7, 14482-14508.

Shafiee-Jood, M. and Cai, X., 2016. Reducing food loss and waste to enhance food security and environmental sustainability. *Environmental science & technology*, 50(16), pp.8432-8443.

Sothe, C., Almeida, C. M. d., Liesenberg, V. & Schimalski, M. B. 2017. Evaluating sentinel-2 and Landsat-8 data to map successional forest stages in a subtropical forest in Southern Brazil. *Remote Sensing, 9, 838.*

Veloso, A., Mermoz, S., Bouvet, A., Le Toan, T., Planells, M., Dejoux, J.-F. & Ceschia, E. 2017. Understanding the temporal behaviour of crops using Sentinel-1 and Sentinel-2-like data for agricultural applications. *Remote Sensing of Environment,* 199**,** 415-426.

Vuolo, F., Neuwirth, M., Immitzer, M., Atzberger, C. & Ng, W.-T. 2018. How much does multitemporal Sentinel-2 data improve crop type classification? *International journal of applied earth observation and geoinformation*, 72, 122-130.

https://earthexplorer.ugs.gov/