Original Research Article

Comparison of crop area estimation for the winter season using the Google Earth Engine (GEE) platform and traditional method of supervisor classification.

Abstract

There is a significant development in the methods and techniques for classifying satellite images during these years. One of the most recently launched methods is using the Google Earth Engine platform for its contribution in providing a huge amount of satellite data with different dates and multiple satellite images sources, with a method of work that is easy to deal with, but it needed to learn the computer language such as Java and Python, which is the method of work depends on writing a script that gives application commands to classify the different satellite image.Therefore, this research aims to apply classification of satellite images using the Google Earth Engine platform for three models Classification and Regression Trees (CART), Support Vector Machine (SVM), and Random Forest (RF), then compare the results calculated from using images classification (Supervised classification: Maximum likelihood) in terms of the degree of accuracy of the classification and matching the results of the classified satellite images in addition. to boundary clarify of the classification methods.This study showed that, using the Google Earth Engine (GEE) platform is considered one of the important methods that facilitated researchers interested in using remote sensing techniques and geographic information systems, due to the availability of this platform of free satellite sources with multiple spatial resolutions, in addition to the presence of a powerful system that collects, processes, and classifies satellite images using different classification models and indexes.But we need to apply these modern techniques at a number of levels, such as at Markas, Governorates, International, and regional levels, to judge their accuracy and suitability for large areas.

Key word:GEE platform, CART, SVM, RF, supervised classification, MLKH, Qewsina.

1- Introduction

The winter crop areas were identified and mapped at the field level and aggregated at the governorate scale in the period from 2013 to 2019 using GEE and RF classifier with temporal satellite data. from Landsat 8, and linked with collected crop data from the United States Department of Agriculture (USDA)-Natural Resources Conservation Service (NRCS). GEE platform provides free, processed satellite imagery and powerful computing power, which processes this data to solve the challenging problem of big data processing for remote sensing. Based on GEE platform, this study was designed to use annual spatiotemporal binary datasets with 30-meter spatial resolution from 2013 to 2019. It then clustered them by study area boundaries using Landsat 8 imagery and data collection. In addition, spectral signature, land cover crop patterns, and linear changes in crop cover NDVI were explored and evaluated in this study **(Ahmed et al. 2023).**The cloud computing platform of GEE and multiple satellite data were used to analyze the long-term and large-scale spatiotemporal scale of wetlands in the study area over a 20-year period (2000 – 2020). They conclude that these platforms are of great importance to environmental planners and ecosystem management specialists because they serve as basic information necessary to develop effective strategies to mitigate the negative impacts of climate change, especially on semi-arid ecosystems such as wetlands **(Gxokwe et al. 2023)**.The accuracy of four machine learning classifiers available on GEE, RF, CART, SVM, and Naïve Bayes (NB), were compared using high-resolution satellite images. Moderate (MODIS) and BRDF modified reflectance data (MCD43A4 V6) and BRDF and Albedo model parameter data (MCD43A1 V6) are used as satellite image sources. They explained that the classifiers were ranked according to accuracy, from highest to lowest: RF, CART, SVM, and NB. Finally, they concluded that this study contributes to the development of crop area mapping and the application of multi-angle monitoring satellites **(Zhen et al., 2023)**. The application of machine learning technology using RF to classify Landsat satellite images in the mountainous terrain of the Himalayas in the western side of the Indian state was presented. The accuracy of the classification depended on the algorithm used and the difference in reflectance behavior of the different classification sections of interest. This requires further study to effectively integrate different factors that will improve the accuracy of map production. Using newer available machine learning techniques and formulating new algorithms to improve plant classification through is another research area that takes advantage of the availability of remote sensing data sources **(Pande et al. 2023).**

Land uses are classified through the RF algorithm method with the help of several indicators implemented to obtain the vegetation area that forms part of the urban forest. They concluded that the results of this study are expected to be recommendations and considerations in policy making for urban forest management and reducing the impact of urban communities **(Ranti et al., 2022).** Some study aims to draw an accurate and early map of crop types in a heterogeneous agricultural area with very small areas surrounding irrigated lands. To conduct this study, the high spatiotemporal resolution of Sentinel-1 and Sentinel-2 satellite images and the RF classifier was applied on GEE platform. They showed that red-edge and shortwave infrared bands could improve crop classification accuracy by 1.72% compared with using only conventional bands (i.e., visible and near-infrared bands). They also concluded that GEE processing was highly efficient in accessing remote sensing data and saving time **(Saad Al-Eman et al., 2022).** GEE platform was used to monitor changes in soil erosion of Ordos region from 2013 to 2021. It would be concluded that, using GEE platform. It is possible to interpret and give a better understanding of the current situation and changes in soil and water loss and identify the origin or least formed landscape through track indicators of historical changes to provide support for restoration and management of environmental protection **(Liu et al., 2022).** Satellite image classification was used applying vector classification algorithms including NNC (Neural Network Classification), MLC (Maximum Likelihood Classification), SVM, MinD (Minimum Distance), and MahD (Mahalanobis Distance)to produce land cover maps through the Google Earth Engine platform. The results of this research indicated acceptable accuracy of NDVI index values for producing natural land cover maps and can be used in preparing these maps for geographical monitoring and achieving sustainable environmental development**. (Muhammad et al. 2022).** Land cover was classified and compared using RF and SVM techniques. The classification result of the land cover map was divided into 5 sections, including forests, agricultural land, grass, soil, and water, and the overall accuracy of the machine learning method was higher than 87%. **(Jarjaldalai et al., 2022).** Two decades (2000 to 2020) medium resolution satellite datasets (MODIS) and (MOD13Q1) were used to analyze the NDVI vegetation trend across Iran. First, an annual per-pixel Difference Vegetation Index (NDVI) dataset was prepared using GEE platform by averaging all available NDVI values over the growing season, and then fed into the PolyTrend algorithm to determine the trend Linear/non-linear. **(Arslan et al. 2022).**

The use of different classification algorithms such as MinD, RF, SVM, CART, and NB (Naive Bayes) were evaluated in an agricultural region of Italy. They conclude that combining S1 and S2 satellite images improves the classification slightly with respect to the classification results of S2 satellite images only. On the other hand, the use of time series significantly improves directed classifications. A future development of the research is to test the advantages brought by the integration of Sentinel-1 and Sentinel-2 time series **(Clemente et al., 2020)**. A number of 2199 Landsat satellite images were used on the GEE platform and the EVI (Enhanced Vegetation Index) and harmonic regression methods were applied to determine the phenological behaviors of land cover sections during leaf shade (LSP) and leaf fall seasons (LFS). They concluded that, GEE cloud platform is a suitable tool to use to access remote sensing of big data at large scale without cost **(Venkatappa et al., 2019).**Exploring the efficiency of using the GEE platform when classifying multi-temporal satellite images with the possibility of applying the platform on a broader scale up to the country level and multiple satellite image sources such as Landsat-8 and Sentinel-2 provides very good performance in terms of enabling access to remote sensing data through the cloud platform and provided pre-processing. However, in terms of classification accuracy, the web-based approach outperformed the SVM, CART and RF classifiers available in GEE platform**(Shelestov et al. 2017)**.

2- Materials, software, and Methods:

2-1 Materials and software's

- The Universal Transferase Mercator (UTM) projection was used for all satellite images, and maps covered the studied areas.
- The administrative borders of the Arab Republic of Egypt include the borders of regions at the level of governorates and centers (source: Organization and Administration Authority 2017).
- Sentinel 2 satellite images dated March 2023.
- Geographical distribution map of field investigation points for winter season 2023.
- The dynamic GPS was used to collect field check data during winter season 2023.
- ERDAS IMAGINE 2022 software was performed for band combination, image subset, visual interpretation, and supervised classification.
- ArcGIS 10.8 software was used for GIS purposes, doing the following: Input data in various formats, data conversion, Overlaying the layers for image interpretation, digitizing, area estimation.

2-2 Method used

- Determine the administrative boundaries of the Markas Qewsinausing the ArcGIS program 10.8. the ship file of Markas Qewsinawas uploaded on GEE platform.
- Selected start and end searchesdate ofSentinel 2satellite images commensurate with the completion of vegetative growth of differentwinter season crops during the period from 1 Marh to 15 march 2023.

- Uploading select Sentinel 2satellite images of winter season in March 2023 and the administrative boundaries of the Markas Qewsinaonto the dynamic global positioning device.
- Collecting observations points during field investigation for different crops types of winter season 2023 covered Markas Qewsinaand recording them on the device in the form of ship files.
- Applying the framework method of the GEE platform for three different classification methods named SVM, CART and RF classifiers.
- Applying the guided classification of supervised classification(Maximum Likelihood classifier)using the ERDAS Imagine 2022 program.
- Estimating the general and specific accuracy degrees for each section for the three classification methods on the GEE platform, as well as for the directed classification method of supervised classificationon IRDAS Imagine 2022.
- Calculating and comparing the areas of classified crops for the three methods using the GEE platform and the directed classification method of supervised classificationusing IRDAS Imagine 2022.

3- Results and discussion

3-1 Location of Markas Qewsina

Markas Qewsina is located in Al Monofia Governorates which is covered 49607.8 Feddans. **Figure (1)** shows the location of Markas Qewsina. It is situated between 31˚ 02´ 59.94˝ E and 31˚ 16´ 00.90˝ E and between 30˚ 27´ 55.95˝ N and 30˚ 39´ 10.56˝ N.

3-2 Crop areas estimation

The cropped area for the winter season 2023 of Markas Qewsinawas calculated using three classification methods on the GEE platform (CART - SVM - RF), which represent modern methods on the Internet, and it was also calculated using traditional methods represented by supervised classification.

3-1-1 Calculate Crop areas estimation using GEE platform

GEE method was used to calculate crop areas estimation by applied classification of three models CART, SVM, and RF. These three models were applied through the GEE web after registration accepted. Jafa script was written for the three models used. The following steps were explained the process of using GEE platform.

- **Step one: identified the boundaries of the Markas Qewsina.**

The shape file of Markas Qewsinaboundary was uploaded on GEE platform. **Figure (2)** show the result of uploaded Markas Qewsinaboundary shape file.

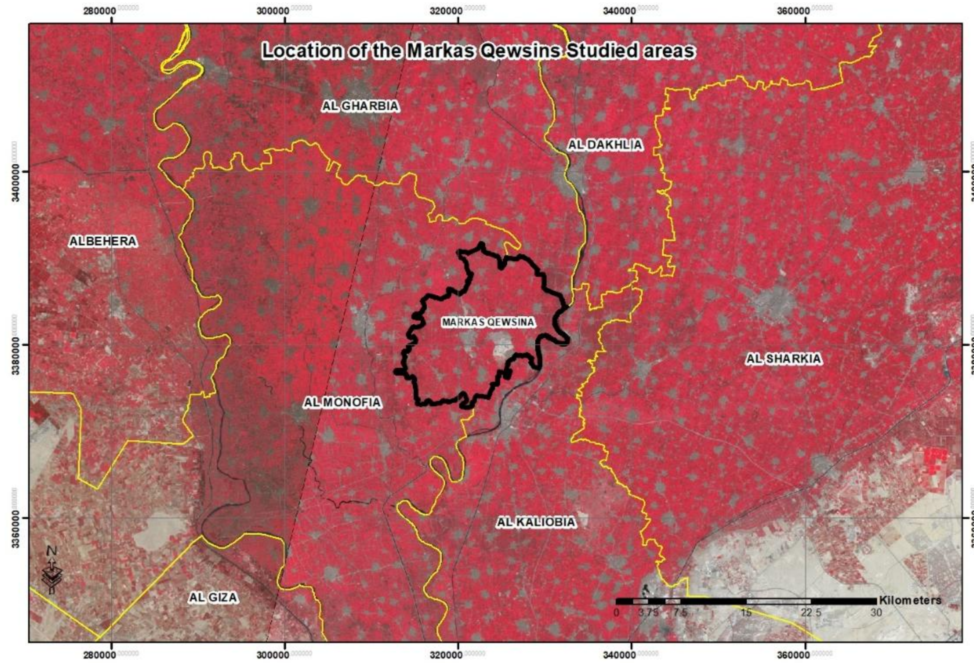

Figure (1): The locations of Markas Qewsina in Monifia Governorate, Egypt.

- **Step Two: identified the satellite images used.**

The satellite images types started date, ended date, and cloud % of Markas Qewsinawas identified using the GEE platform rules. **Figure (3)** show the result of satellite images used to cover Markas Qewsinaarea.

- **Step Three: uploaded Training points**

The shape file of field training points included all field observation were uploaded in the system of GEE platform and identified the 70% random point for area classification and 30% random points for classification validation. **Figure (4)** show the result of uploaded all training points of Markas Qewsina. A total of 962 field observations were used, which were randomly divided into 672 points (70% of the total points) to conduct the classification process and 290 points (30% of the total points) to conduct the process of estimating the degree of classification accuracy usedGEE platform.

- **Step Four: applied the three model and calculate the accuracy**

The three models CART, SVM, and RF were applied and calculated the Confusion Matrix, Total Accuracy, Producer's accuracy, and Kappa statistic of the applied three models. Figure (6, 7, and 8) show the results of applied the three models on GEE platform.The total Accuracy of CART, SVM and, RF classification methods were 0.900, 0.907, and 0.897 respectively. The producer's accuracy of CART classified classes was Wheat (0.877), Alfalfa (0.786), Other Crops (0.914), Bare Soils (0.556), Urban (0.9530), Water Bodies (1.000). The producer's accuracy of SVM classified classes was Wheat (0.862), Alfalfa (0.857), Other Crops (0.948), Bare Soils (0.611), Urban (0.963), Water Bodies (0.926).The producer's accuracy of RF classified classes was Wheat (0.877), Alfalfa (0.786), Other Crops (0.897), Bare Soils (0.722), Urban (0.944), Water Bodies (0.926).The Kappa statistic of CART, SVM, and RF classified methods were 0.868, 0.876, and 0.864respectively.

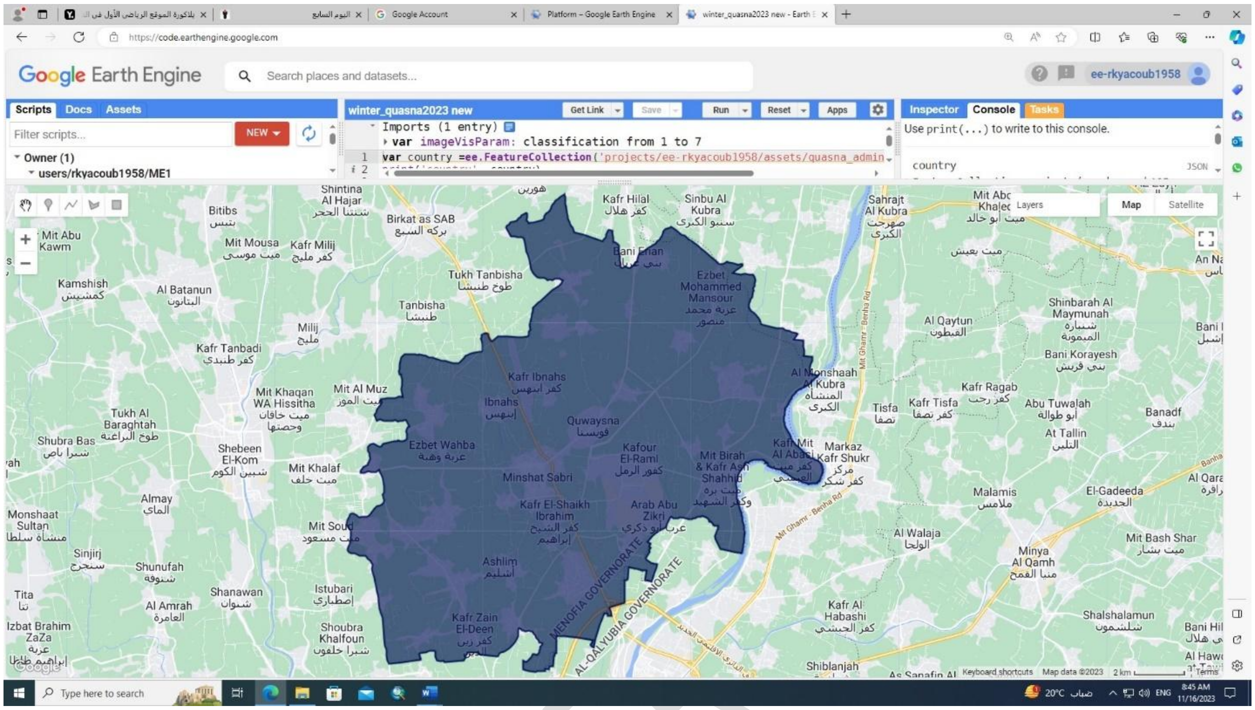

Figure (2) the result of uploaded Markas Qewsinaboundary shape file.

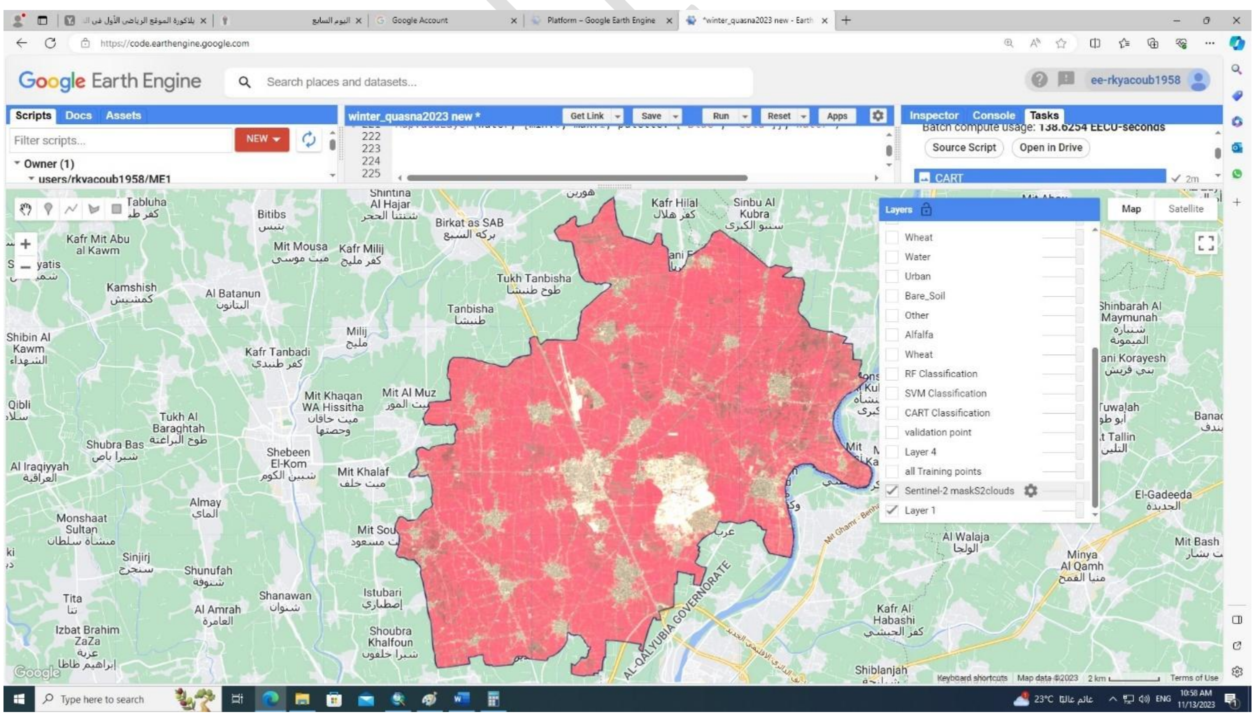

Figure (3) the result of satellite images used for Markas Qewsina

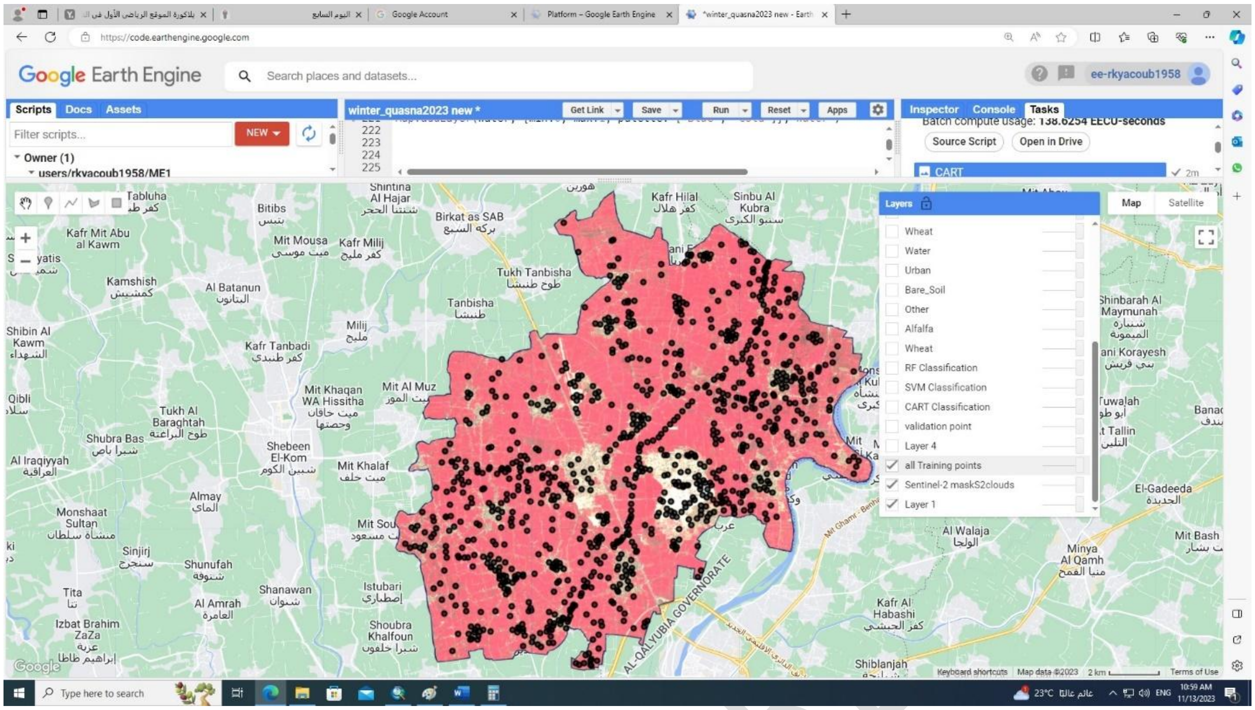

Figure (4) the result of uploaded All training points of Markas Qewsina.

- **Step five: Export the areas classes of the three model and calculate areas**

The results of the three models CART, SVM, and RF were exported and calculated. **Figure (5, 6, & 7)** show the results of applied GEE platformusing the three models' classification. **Figure (8, 9, &10)** show the results of exported classes areas using the three models' classification. The exported results of the three models were treated using the Arc-GIS software environment and converted from raster format to shape format then reclassified based on its code.Also, the area calculation of each class was applied to the three models.**Tables 1&2** show the exported areas classes and Object areas of 6 classes according classified CART, SVM, and RF classified in Feddan.

Table (1) the exported areas classes using CART, SVM, and RFclassified using Arc-GIS software.

Classes	CART model		SVMmodel		RF model	
	Area in		Area in		Area in	
	Feddan	%	Feddan	%	Feddan	%
Wheat	23765.8	47.91	22748.6	45.86	25491.5	51.39
Alfalfa	4754.5	9.58	4528	9.13	4487.4	9.05
Other Crops	10441.2	21.05	13210.3	26.63	9298.1	18.74
Bare Soils	390.9	0.79	610.8	1.23	769.3	1.55
Urban	9834.1	19.82	8238.7	16.61	9039.8	18.22
Water Bodies	421.3	0.85	271.3	0.55	521.7	1.05
Total area	49607.8	100	49607.8	100	49607.8	100

Table (2) Object areas classes according CART, SVM, and RFclassified in Feddan using GEE platform

Classes	CART model		SVMmodel		RF model	
	Area in		Area in		Area in	
	Feddan	%	Feddan	%	Feddan	%
Wheat	22254	44.86	22466	45.29	23439	47.25
Alfalfa	5442	10.97	4752	9.58	5000	10.08
Other Crops	11341	22.86	13251	26.71	10511	21.19
Bare Soils	613	1.24	620	1.25	399	0.8
Urban	9480	19.11	8230	16.59	9808	19.77
Water Bodies	477	0.96	290	0.58	452	0.91
Total area	49607	100	49609	100	49609	100

The results of calculated areas classes using conversion tools in Arc-GIS for each classification method on GEEplatform show some differences with the results of using calculated areas direct on GEE. The results in these Tables show that, only area calculation using SVM classified were almost same in the two-calculation method that then the other two-classified method (CART, and RF). From that there is effect of classification accuracy on the results of areas calculation.

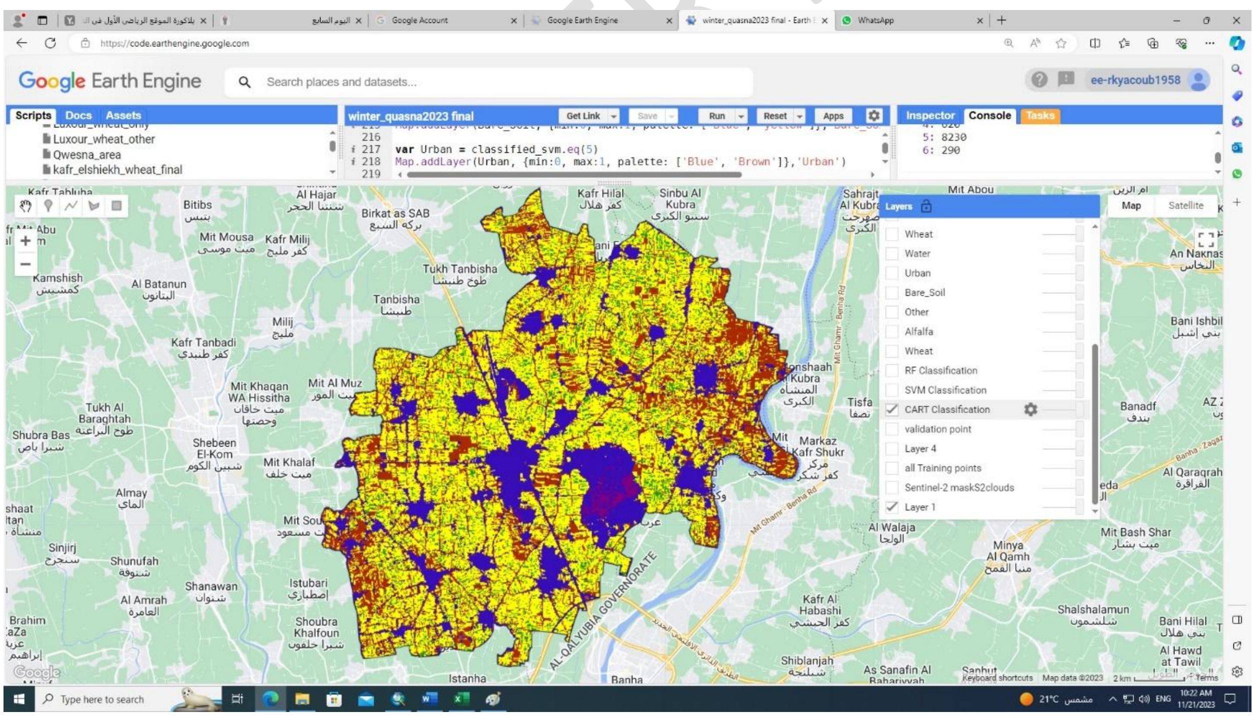

Figure (5) the results of applied CART using GEE platform.

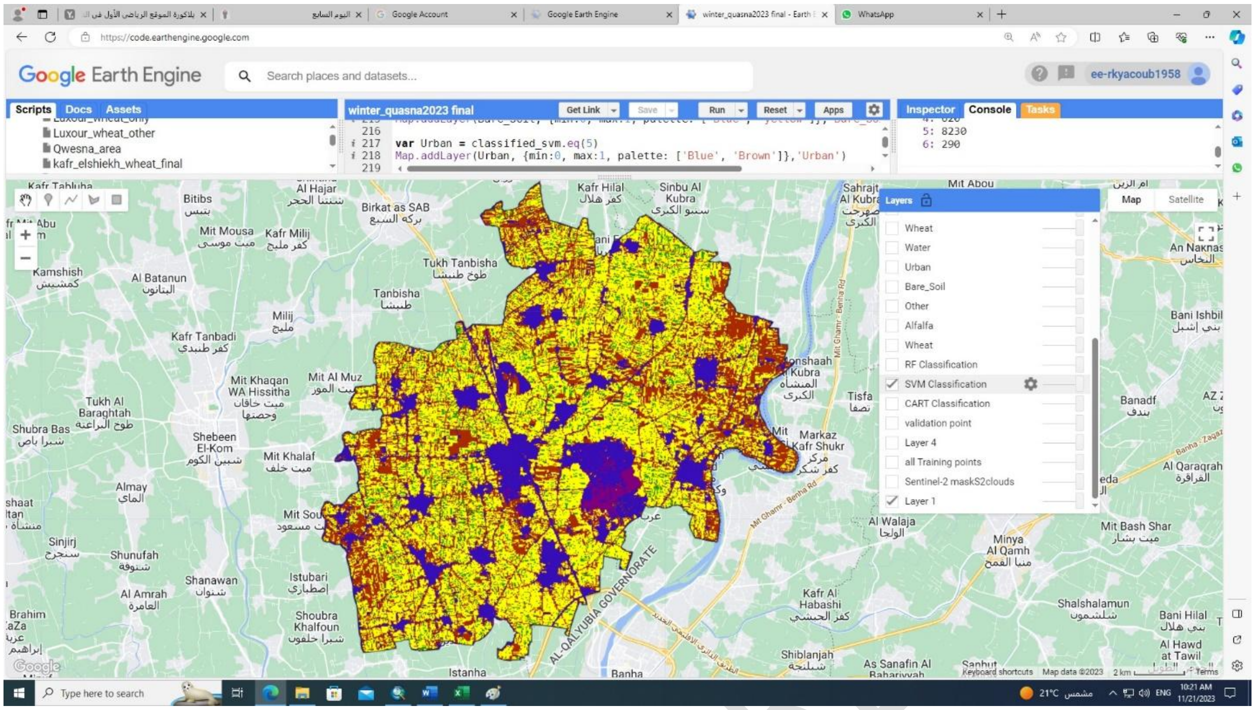

Figure (6) the results of applied SVM using GEE platform.

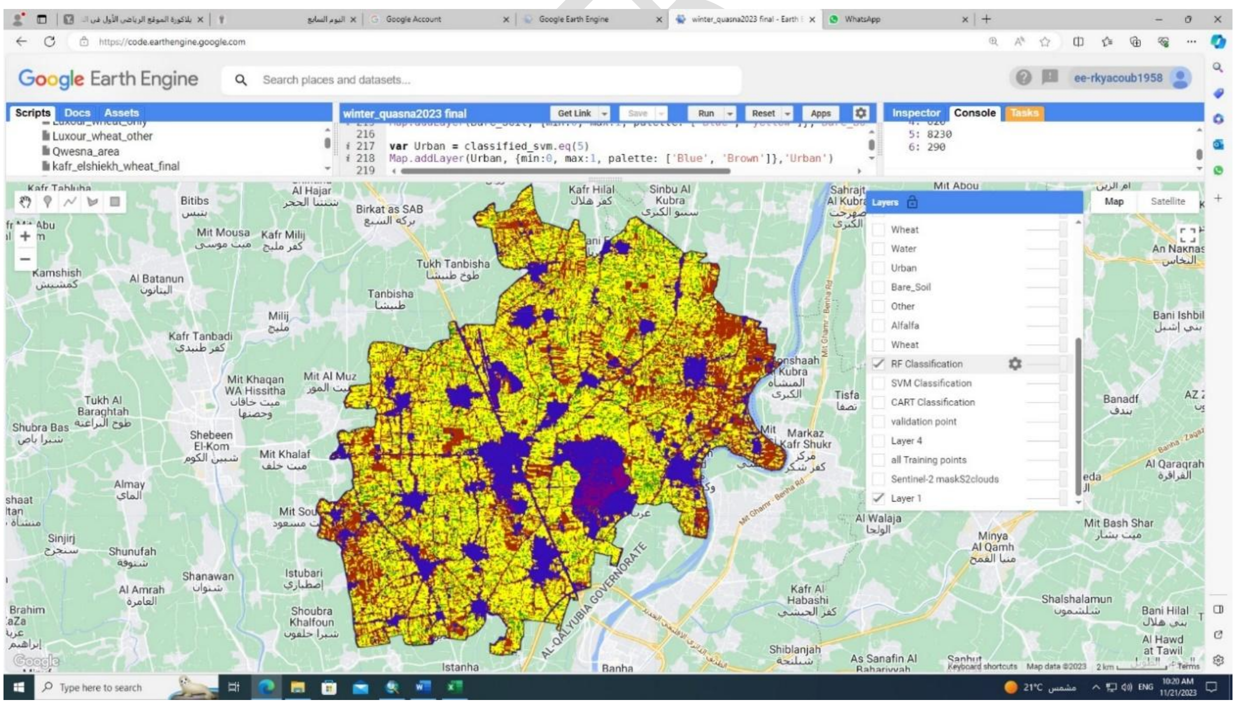

Figure (7) the results of applied RF using Gee platform.

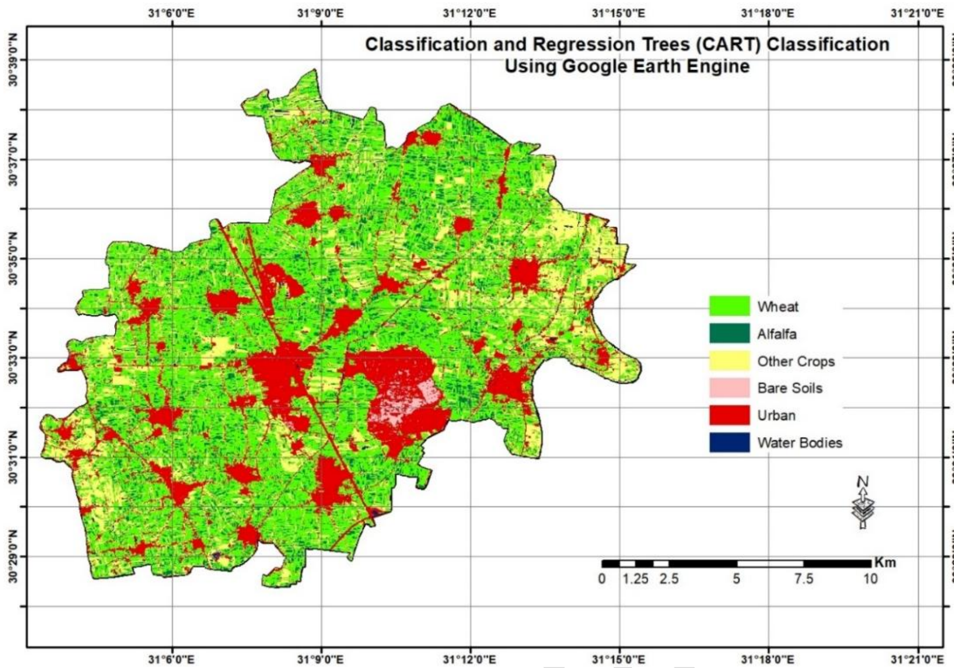

Figure (8) the results of exported the CARTclassifier.

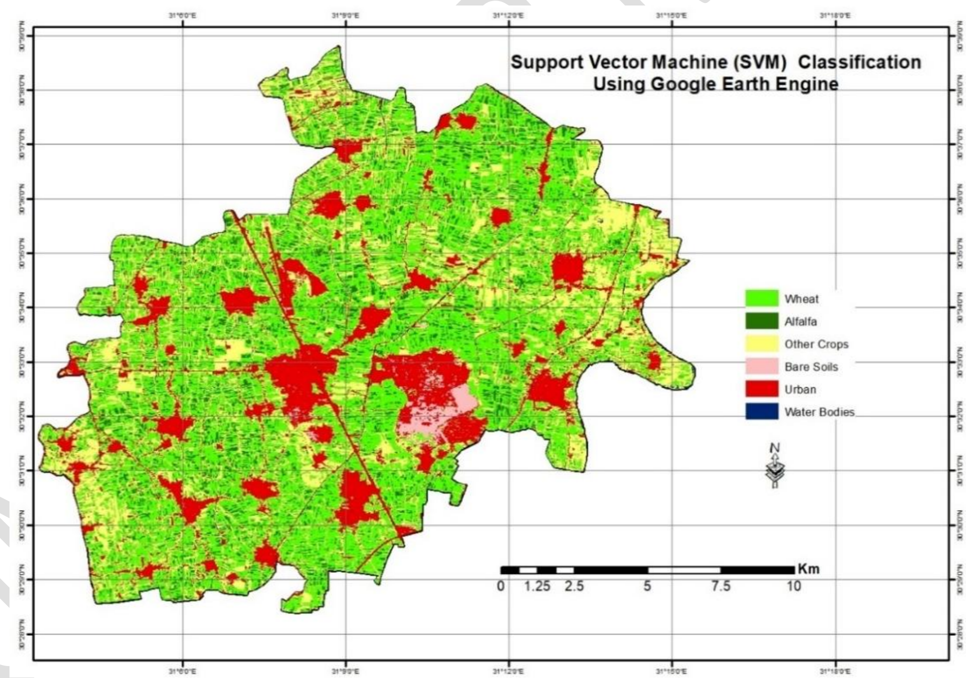

Figure (9) the results of export the SVM classifier.

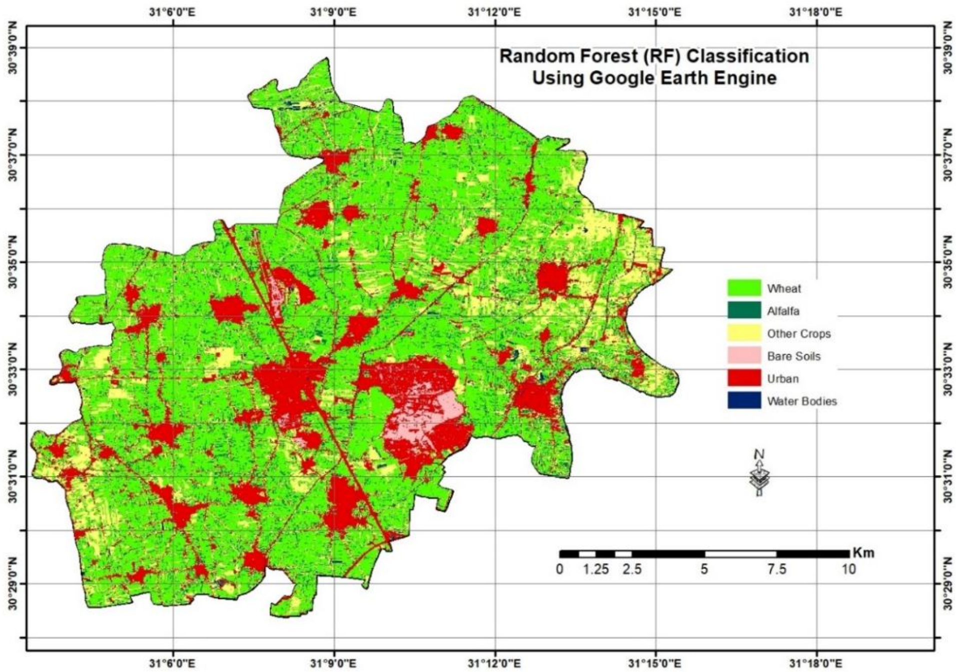

Figure (10) the results of export the RFclassifier.

3-1-2 Calculate Crop areas estimation using supervised classification(Maximum Likelihood -MLKH)

The Sentinel 2 image was downloaded from https://dataspace.copernicus.eu. And the acquisition date is 4 March 2023. Through the ERDAS Imagine 2022 program, The satellite image was prepared for the classification process using the supervised classification method, where the spectral fingerprints were randomly divided in a ratio of 30% and 70% for each field data collected. The percentage of 70% from collected field work (672 points) were used in the supervised classification process and the other 30% of the collected field data (290 points) were used to calculate the Accuracy Assessment. This random field data collect (70% of the collected field data) were converted from points to areas (Point To Polygon operation) using the Buffering technique and setting Linear unit = 1 m reduce. The created 672 training samples areas were identified for 6 predefined land cover classes, named Wheat, Alfalfa, Other Crops, Bare Soils, Urban areas, and Water Bodies. The standard deviation of extracted signature of each class ranged from 0.8 to 10. The Maximum Likelihood method of supervised classificationimage (MLKH) was applied to produce land use maps. **Figure (11)** shows MLKH land use map for 4 March 2023. **Table (3)** show the calculated areas of MLKH land use maps classes for 4 March 2023.The accuracy assessment of the supervise classification method was evaluated using Arc/GIS version 10.8, where 30% of the collected field data were used in steps as follows:

- Extract Input raster values of Markas Qewsinasupervise2023 to Points
- Frequency Field of classify raster value
- Pivot tableof Fieldclassify raster value frequency
- Convert Output table to Excel Sheet.
- Added the land use classes using classes code.
- Calculated the sum of random points in horizontal and vertical direction.

- Calculated the Kappa values using web http://vassarstats.net/kappa.html.

Table (4) show the results of applied the accuracy assessment of the supervise classification method and kappa calculated.

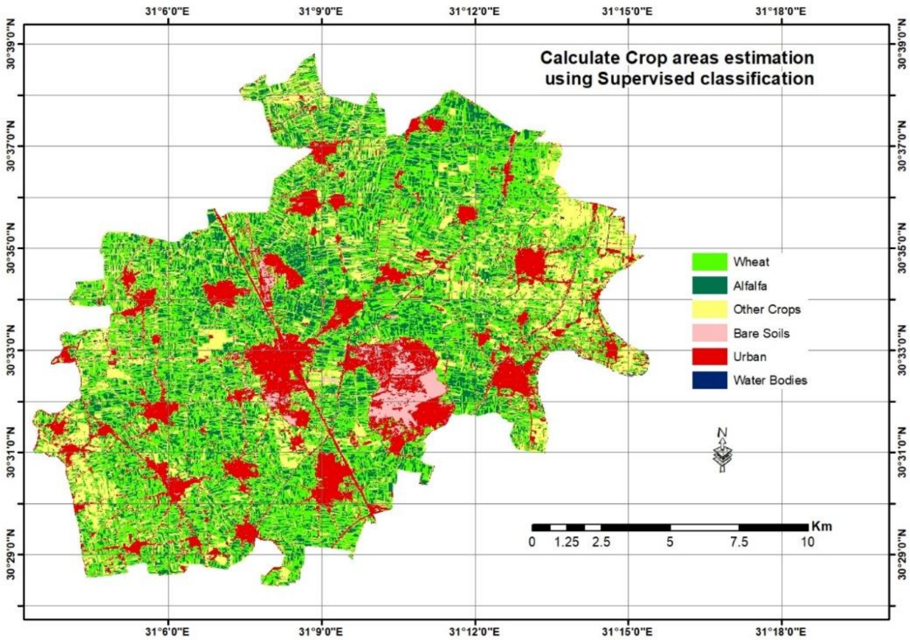

Figure (11) MLKH land use map for 4 March 2023

Table (3) calculated areas of MLKH land use maps classes for 4 March 2023.

classes	Area in	
	Feddan	%
Wheat	19109.57	38.52
Alfalfa	9219.9	18.59
Other Crops	10661.32	21.49
Bare Soils	2016.597	4.07
Urban	8317.275	16.77
Water Bodies	283.2079	0.57
Total area	49607.9	100.00

1- Conclusions

This study showed that, using the GEE platform is considered one of the important methods that facilitated researchers interested in using RS and GIStechniques, due to the availability of this platform of free satellite sources with multiple spatial resolutions, in addition to the presence of a powerful system that collects, processes, and classifies satellite images using different classification models and indexes.In general, the degree of accuracy for both the three classification methods used on the GEE platform and also the result when using supervised classification is considered very close and ranged between 0.879 to 0.908. The highest value was for the SVM classification method and the lowest value was for the supervised classification method, although the difference between the highest value and the lowest value was Only 0.029.

Despite the ease of classification methods on the GEE platform, the user of these techniques needs to be familiar with the Java and Python languages, which are relied upon to implement commands to specify the region and the selected satellite images. Any error in the writing text,the results give an error and failure to complete the implementation of the classification methods.

However, the result of the change in classification results when using the three classification methods puts researchers in making an important decision to choose the optimal method to use based on the degree of classification accuracy, which may be inappropriate for the geographical distribution of the different classification sections. It is also difficult to compare the classification methods found on the GEE platform with traditional methods such as supervised classification, which appeared in a decrease in the areas of the classification sections. For example, the areas of the wheat crop were calculated using the GEE platform, and the percentage was greater than 44% in the three classification methods, while it did not reach 39% when using supervised classification.

Table (4) the results of applied the accuracy assessment and kappa calculated

Classes	Wheat	Alfalfa	Other crops	Bare soils	Urban	Water Bodies	points	User Accuracy
Wheat	**39**	7	4	0	0	0	50	0.780
Alfalfa	4	**23**	1	0	0	0	28	0.821
Other crops	0	0	**69**	0	1	3	73	0.945
Bare soils	0	0	0	**10**	5	0	15	0.667
Urban	0	0	0	9	**97**	0	106	0.915
Water Bodies	0	0	1	0	0	**17**	18	0.944
points	43	30	75	19	103	20	**290**	
Producers Accuracy	0.907	0.767	0.920	0.526	0.942	0.850		
Total accurate point								**255**
Overall accuracy								0.879
Kappa								0.842

2- References

A Ranti, R Asy'Ari and T H Ameiliani (2022), "Detection of Urban Forest Change in Jabodetabek Megacity Using Sentinel 2 and Landsat 8 Imagery Through Google Earth Engine Cloud Computing Platform", IOP Conf. Series: Earth and Environmental Science 959 (2022) 012028 IOP Publishing doi:10.1088/1755-1315/959/1/012028.

Arsalan Ghorbanian, Ali Mohammadzadeh and Sadegh Jamali (2022), "Linear and Non-Linear Vegetation Trend Analysis throughout Iran Using Two Decades of MODIS NDVI Imagery", Remote Sensing. 2022, 14, 3683. https://doi.org/10.3390/rs14153683. https://www.mdpi.com/journal/remotesensing.

J. P. Clemente, G. Fontanelli, G. G. Ovando, Y. L. B. Roa, A. Lapini, and E. Santi (2020), "GOOGLE EARTH ENGINE: APPLICATION OF ALGORITHMS FOR REMOTE SENSING OF CROPS IN TUSCANY (ITALY)", The International Archives of the

Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLII-3/W12-2020, 2020 2020 IEEE Latin American GRSS & ISPRS Remote Sensing Conference (LAGIRS 2020), 22–26 March 2020, Santiago, Chile

Jargaldalai Enkhtuya, Amarsaikhan Damdinsuren, Bilguun Ulziibat, and Munkh-Erdene Altangerel (2022), "Land cover classification using machine-learning method and vegetation indices", Mongolian Journal of Geography and Geoecology – MJGG, Volume 43. 2022.

Liu, Y.; Liu, J.; Zheng, Y.; Kang, Y.; Ma, S.; Zhou, J. (2022), "Tracking Changing Evidence of Water Erosion in Ordos Plateau, China Using the Google Earth Engine", Land 2022, 11, 2309. https://doi.org/10.3390/land11122309 https://www.mdpi.com/journal/land

Manjunatha Venkatappa, Nophea Sasaki, Rajendra Prasad Shrestha, Nitin Kumar Tripathi and Hwan-Ok Ma (2019), "Determination of Vegetation Thresholds for Assessing Land Use and Land Use Changes in Cambodia using the Google Earth Engine Cloud-Computing Platform", Remote Sens. 2019, 11, 1514; doi:10.3390/rs11131514

Mohammad Mansour moghaddam, Iman Rousta, and Hamid Reza Ghafarian (2022), "Evaluation of the classification accuracy of NDVI index in the preparation of land cover map", DESERT, 27 (2), DOI: 10.22059/jdesert.2022.90834.

Raj Singh, Arun Pratap Mishra, Manoj Kumar, and Chaitanya B. Pande (2023), "Classification of Vegetation Types in the Mountainous Terrain Using Random Forest Machine Learning Technique", Climate Change Impacts on Natural Resources, Ecosystems and Agricultural Systems, Springer Climate, https://doi.org/10.1007/978-3-031-19059-9_27

Saad El Imanni, H.; El Harti, A.; Hssaisoune, M.; Velastegui-Montoya, A.; Elbouzidi, A.; Addi, M.; El Iysaouy, L.; El Hachimi, J. (2022), "Rapid and Automated Approach for Early Crop Mapping Using Sentinel-1 and Sentinel-2 on Google Earth Engine; A Case of a Highly Heterogeneous and Fragmented Agricultural Region", J. Imaging 2022, 8, 316. https:// doi.org/10.3390/jimaging8120316.

Shelestov A, Lavreniuk M, Kussul N, Novikov A and Skakun S (2017), "Exploring Google Earth Engine Platform for Big Data Processing: Classification of Multi-Temporal Satellite Imagery for Crop Mapping", Front. Earth Sci., 2017, pp: 5:17. doi: 10.3389/feart.2017.00017

Siyamthanda Gxokwe, Timothy Dube, and Dominic Mazvimavi (2023), "An assessment of long-term and large-scale wetlands change dynamics in the Limpopo transboundary river basin using cloud-based Earth observation data", Wetlands Ecol Manage https://doi.org/10.1007/s11273-023-09963-y

www.mdpi.com/journal/remotesensing.

Zhen, Z.; Chen, S.; Yin, T.; Gastellu-Etchegorry, J.-P. (2023), "Improving Crop Mapping by Using Bidirectional Reflectance Distribution Function (BRDF) Signatures with Google Earth Engine.", Remote Sens. 2023, 15, 2761. https://doi.org/10.3390/rs15112761

Zobaer Ahmed, Lawton Nalley, Kristofor Brye, V. Steven Green, Michael Popp, Aaron M. Shew, Lawson Connor (2023), "Winter-time cover crop identification: A remote sensing-based methodological framework for new and rapid data generation", International Journal of Applied Earth Observation and Geoinformation 125 (2023) 103564.