CROPLAND MAPPING USING MULTI-SENSOR TIME SERIES DATA

Maryam Salehi
msalehi@mail.kntu.ac.ir

Abstract *– This study demonstrates an approach for cropland mapping by fusing information from Sentinel-1, Sentinel-2 and Landsat-8 time series data. Various bands and spectral indices are derived from the optical and radar datasets, capturing different characteristics of vegetation phenology through the growing season. Machine learning models utilize the multi-temporal data and indices as their input to identify cropland areas. The methodology is evaluated over several study regions with diverse agricultural practices. Results indicate that the integration of multi-sensor time series data improves cropland classification accuracy compared to using individual sensors*.

Keywords – Cropland, Landsat-8, machine learning, multi-sensor, Sentinel-1, Sentinel-2, spectral indices.

1. INTRODUCTION

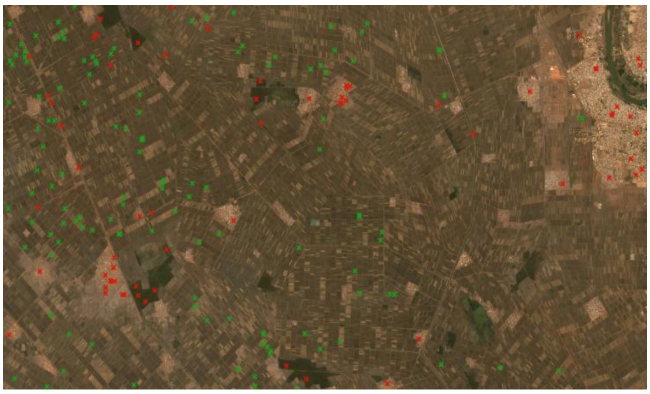
Cropland mapping and monitoring are essential for agricultural management, land-use planning, and food security assessment. The accurate and timely identification of cropland areas is vital for making informed decisions, optimizing resource allocation, and ensuring sustainable land use. In recent years, the advancement of Earth observation satellites, such as Sentinel-1 (S1), Sentinel-2 (S2), and Landsat-8 (L8), has opened up new possibilities for monitoring cropland on a large scale. These sensors provide a wealth of information about the Earth's surface, including optical and radar data, which can be harnessed to monitor changes in cropland over time.

However, each sensor has its limitations, and individual sensors may not capture all the nuances of cropland characteristics. For instance, optical sensors like Sentinel-2 are excellent at providing high-resolution, multi-spectral data, but they are limited by cloud cover and cannot capture information during nighttime. On the other hand, radar sensors like Sentinel-1 are unaffected by weather conditions and can acquire data day and night, but their information is primarily related to surface roughness and moisture content. Landsat-8, with its moderate spatial resolution and multi-spectral capabilities, can provide complementary data to both Sentinel-1 and Sentinel-2.

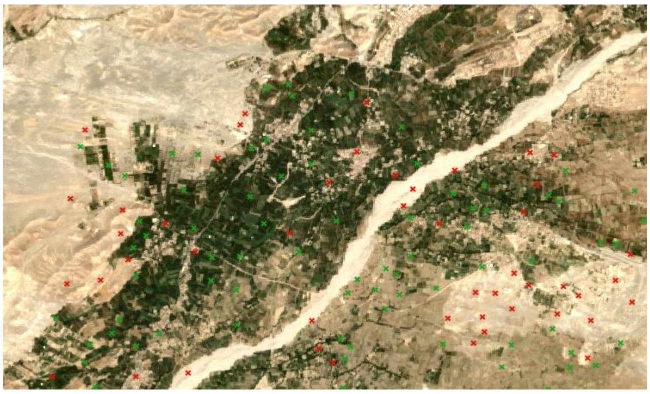
To overcome the limitations of individual sensors and improve the accuracy of cropland mapping and monitoring, we extract various bands and spectral indices from Sentinel-1, Sentinel-2, and Landsat-8 time series. These indices are then employed as input features for machine learning models, enhancing their effectiveness in distinguishing cropland areas.

2. STUDY AREA

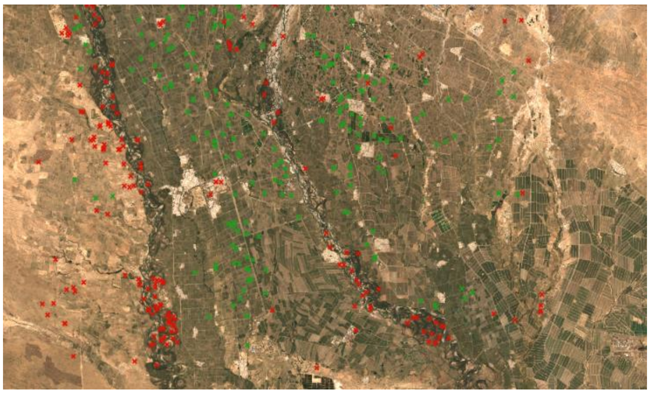
The study encompasses three distinct regions: Sudan, Afghanistan, and Iran, as illustrated in Fig. 1. Ground truth samples were collected during specific time frames, from July 2019 to June 2020 for Sudan and Iran, and in April 2022 for Afghanistan.



(a)



(b)



(c)

Fig. 1 – **Study areas**: (a) Sudan, (b) Afghanistan, and (c) Iran. Green crosses indicate crop areas, while the red crosses represent non-crop areas.

The RGB image (R: B4, G: B3, B: B2) displayed in above figure is a mean image obtained from the Sentinel-2 dataset over specified time frames.

3. METHODOLOGY

3.1 Data Preprocessing

The first step in our methodology entails acquiring and preprocessing data from Sentinel-1, Sentinel-2, and Landsat-8. This data preprocessing phase encompasses several key tasks, including the removal of cloud cover from Sentinel-2 SR (surface reflectance) HARMONIZED and Landsat-8 SR images, selecting VH and VV bands, and calculating the mean values of VH for both ascending and descending orbits, as well as the mean value of VV for the combined ascending and descending orbits of Sentinel-1 GRD images.

3.2 Spectral Index Extraction

Various spectral indices were derived from the optical datasets. These indices include, but are not limited to, the Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), and Soil-Adjusted Vegetation Index (SAVI). These indices capture different aspects of vegetation growth and health, such as leaf area, chlorophyll content, and moisture conditions. Sources of information for these spectral indices can be found in references [1, 2].

3.3 Machine Learning Models

Machine learning models, including XGBoost, AdaBoost, and CatBoost, were individually trained for each country using the multi-temporal bands and indices as input features. These models were designed to differentiate between cropland and non-cropland areas by learning from the distinct patterns presented in the time series data. Notably, the CatBoost Classifier exhibited the highest level of accuracy. Model performance was evaluated based on accuracy and log loss.

4. RESULTS

To evaluate the effectiveness of our approach, we applied the methodology to three study regions. Based on the obtained results presented in Tables 1-4, we can make the following observations about using multi-sensor time series data for cropland mapping:

- Using Sentinel-1 (S1) data alone provides decent accuracy for cropland classification, with accuracy scores ranging from 83-95% across the three study regions (Table 1). However, S1 has higher log loss scores, indicating less certainty in predictions.

- Using only Sentinel-2 (S2) data improves accuracy and reduces log loss compared to using S1 alone, with accuracy ranging from 85% to 99% and a lower log loss (Table 2). This improvement is likely attributed to the additional spectral bands and indices from S2, which offer more detailed information about plant properties such as chlorophyll content, moisture levels, leaf area, and more.

- Fusing S1 and S2 data further boosts accuracy for Afghanistan to 87%, while maintaining high accuracy for Sudan and Iran (Table 3). Log loss is also reduced for Iran. This demonstrates the value of combining optical and radar data.

- Adding Landsat-8 to the S1+S2 fusion does not yield further improvement (Table 4). The likely reasons that Landsat-8 does not enhance results are as follows:
  - Landsat-8 possesses a coarser spatial resolution (30m) in comparison to Sentinel-2 (10m), thus failing to contribute additional spatial detail.
  - Its spectral bands are similar to those provided by Sentinel-2, leading to redundancy in spectral information.

Table 1 – Performance of the best model (70% Training, 30% Validation) using S1

Study area	Accuracy & Loss
Sudan	0.92, 2.88
Afghanistan	0.83, 6.25
Iran	0.95, 1.68

Table 2 – Performance of the best model (70% Training, 30% Validation) using S2

Study area	Accuracy & Loss
Sudan	0.99, 0.24
Afghanistan	0.85, 5.29
Iran	0.96, 1.2

Table 3 – Performance of the best model (70% Training, 30% Validation) using S1 and S2

Study area	Accuracy & Loss
Sudan	0.99, 0.24
Afghanistan	0.87, 4.57
Iran	0.97, 0.96

Table 4 – Performance of the best model (70% Training, 30% Validation) using S1, S2 and L8

Study area	Accuracy & Loss
Sudan	0.99, 0.24
Afghanistan	0.86, 5.05
Iran	0.97, 0.96

5. CONCLUSION

This study underscores the effectiveness of Sentinel-2 in improving cropland classification accuracy, particularly when combined with Sentinel-1 data. While Landsat-8 did not yield additional benefits, the integration of optical and radar data remains a promising approach for accurate cropland mapping and monitoring in diverse regions. These findings offer valuable insights for land use and agriculture monitoring applications, highlighting the importance of selecting the most suitable satellite data sources for specific research objectives.

REFERENCES

[1] https://www.indexdatabase.de/db/is.php?sensor_id=168
[2] https://www.indexdatabase.de/db/is.php?sensor_id=96