CROPLAND MAPPING: COMPARATIVE ANALYSIS BETWEEN WEIGHTED ENSEMBLE MODELS AND TIME SERIES MODELS

Mohammad Alasawedah[1]
1Eurac Research, Institute for Earth Observation, Viale Druso 1, 39100 Bolzano, Italy

1. INTRODUCTION

Cropland mapping is essential to ensure food security and manage natural resources [1, 2]. Satellite data has increasingly gained popularity for producing these maps at various scales [3, 4, 5]. Many optical Earth Observation (EO) data are utilized for cropland mapping, [6] used Sentinel-2 imagery to map croplands over central north of Italy. However, the use of optical EO data for mapping croplands is constrained by cloud contamination. To overcome this limitation, synthetic aperture radar (SAR) has been used due to its ability to penetrate clouds and not affected by the atmospheric conditions [7, 8]. Notably, other studies have explored the combined use of both optical and SAR data [9].

The evaluation of various machine learning (ML) models, incorporating various EO data to map croplands, demonstrates promising results. A wide range of ML models has been employed, including traditional ones like Random Forest (RF) and Maximum Likelihood classifiers [9], as well as advanced time series models such as the Recurrent-Convolutional Neural Network (R-CNN) [6].

This study presents a comparative analysis between a weighted ensemble and a time-series models to get insights into their respective performance in mapping croplands. A weighted ensemble incorporating Light Gradient Boosting Machine (LGBM) and Categorical Boosting (CatBoost), and a time-series model represented by the InceptionTime model, an ensemble of Convolutional Neural Networks (CNNs), designed to recognize patterns within time series datasets.

2. STUDY AREA AND DATA

Area of interests are shown in figure 1, figure 2, and figure3, representing three distinct regions: Afghanistan, Iran, and Sudan. Cropland extent data for these regions, at a 10-meter spatial resolution, was complemented with ground truth samples collected during specific time frames. Table 1 shows the area of intrests and their time frames.

Table 1 – Area of intrests and their time frames.

Region	Time Interval
Afghanistan	April 2022
Iran	July 2019 ~ June 2020
Sudan	July 2019 ~ June 2020

This work exploited the potential of Google Eart Engine (GEE) to obtain the data from three different sources Sentinel-1, Sentinel-2, and soil moisture. Monthly values were extracted within their respective time intervals. However, for Afghanistan, we extracted monthly mean values for the past three months (January, February, March) to enable near-real-time assessment, as a one-month interval would have limited cropland mapping accuracy.

The extracted data from these sources is as follows:

1. Sentinel-1: This collection includes the S1 Ground Range Detected (GRD) scenes, calibrated and ortho-corrected. Monthly mean values have been extracted from polarization bands, including:
   a. VV: Single co-polarization, vertical transmit/vertical receive.
   b. HH: Single co-polarization, horizontal transmit/horizontal receive.

The combination of these two bands led to the creation of additional features, such as the Radar Vegetation Index (RVI), which is a product of the two polarization bands, among others. The formulas for these features can be found in the associated code.

2. Sentinel-2: this collection includes Harmonized Sentinel-2 MSI, level-2A. Information from various spectral bands have been extracted, with each represented by the monthly mean value. the following bands are included:
   a. B2: Blue band.
   b. B3: Green band.

c. B4: Red band.

d. B5: Red Edge 1 band .

e. B6: Red Edge 2 band.

f. B7: Red Edge 3 band.

g. B8: NIR band.

h. B8A: Red Edge 4 band.

i. B11: SWIR 1 band.

j. B12: SWIR 2 band.

The dataset was expanded by incorporating additional features such as the Normalized Difference Vegetation Index (NDVI), Bare Soil Index (BSI), Green Normalized Difference Vegetation Index (GNDVI), and others. The associated code provides the formulas for these features.

3. Soil Moisture: this collection includes the SMAP Level-4 (L4) Soil Moisture product. Information from the following bands have been extracted:

a. sm_surface: Top layer soil moisture (0-5 cm).

b. sm_rootzone: Root zone soil moisture (0-100 cm).

c. sm_surface_wetness: Top layer soil wetness (0-5 cm;wetness units).

d. sm_rootzone_wetness: Root zone soil wetness (0-100 cm;wetness units).

e. surface_temp: Mean land surface temperature .

f. land_evapotranspiration_flux: Evapotranspiration from land

g. vegetation_greenness_fraction: Vegetation "greenness" or fraction of transpiring leaves averaged over the land area.

h. leaf_area_index: Vegetation leaf area index.

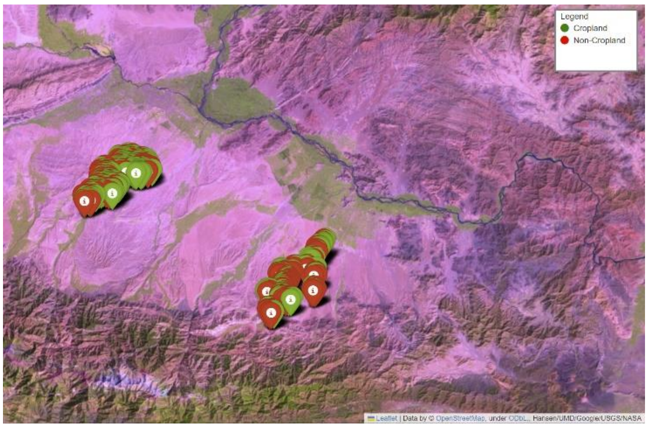

Fig1. – Area of study: Afghanistan, coverd by landsat mosaic 2021.

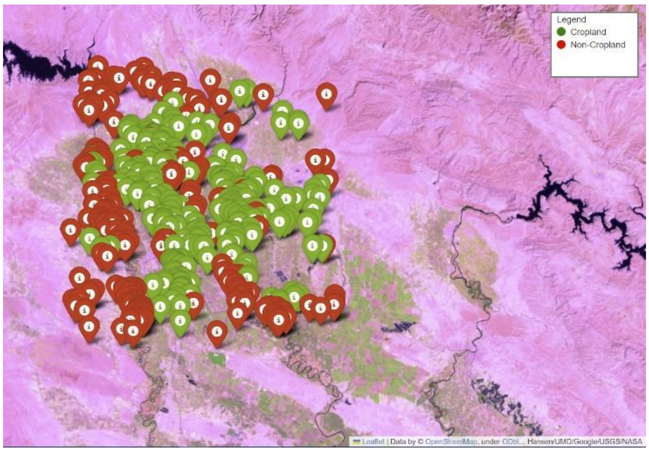

Fig2. – Area of study: Iran, coverd by landsat mosaic 2021.

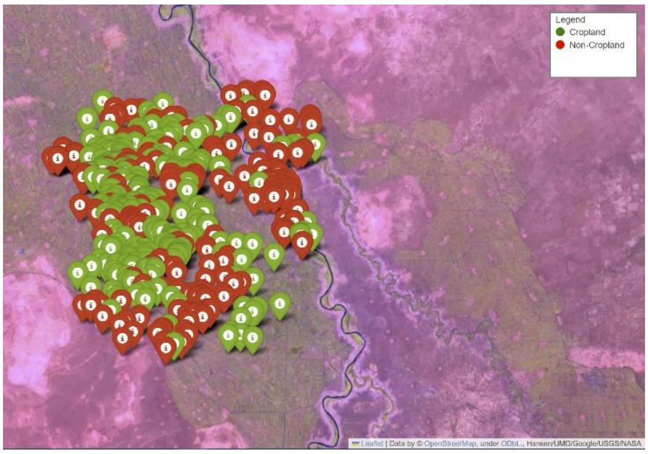

Fig3. – Area of study: Sudan, coverd by landsat mosaic 2021.

The data needed to prepare in a way needed for both methods tree-based models and timeseries model. Timeseries model requires a 3-d array with this format (samples, feature, time) to have (500, 80, 4) for Afghanistan and (500, 80, 13). In contrast, tree-based models required 2-d array with this format (sample, feature) to have (500, 320) for Afghanistan and (500, 1040) for Sudan and Iran. Figure 4 shows the timeseries plots for NDVI, RVI, and soil moisture for samples from Iran, Afghanistan, and Sudan.

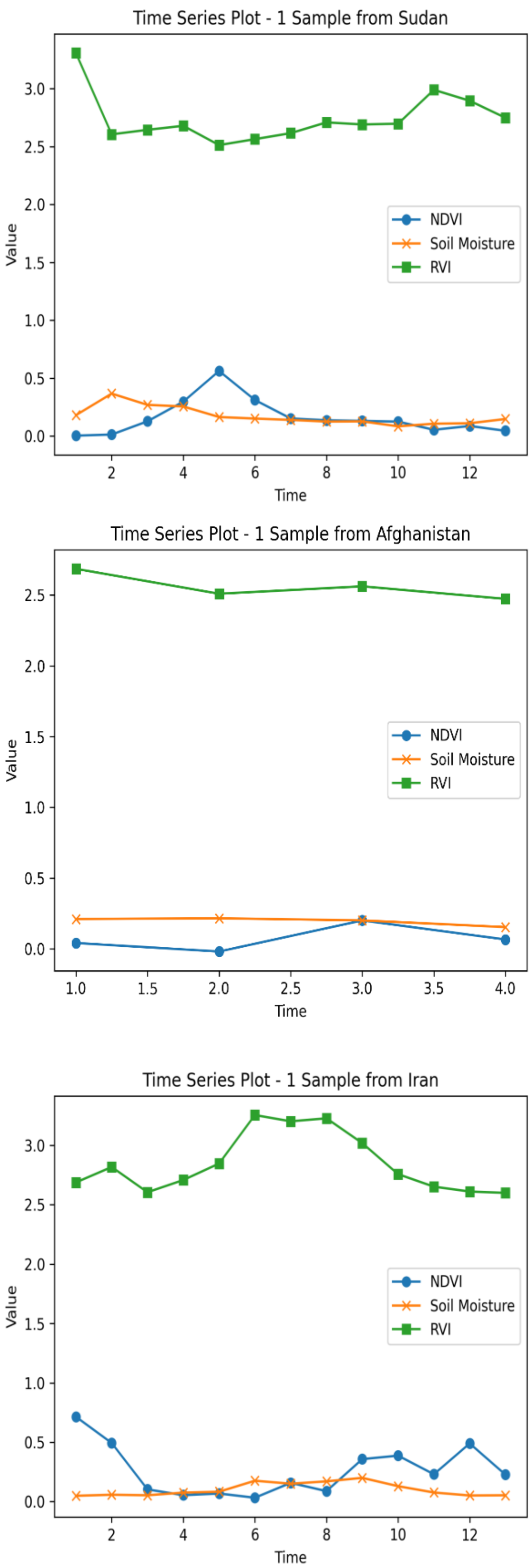

Fig4. – NDVI, Soil Moisture, and RVI values over time for samples from different regions Iran, Afghanistan, and Sudan.

3. RESULTS

For comparison purpose, two different methods were utilized: a time-series model using the InceptionTime model and weighted ensemble models using LGBM and CatBoost. These methods were applied individually to each region for local cross-validation comparison. Subsequently, we combined the test predictions from all three regions in one file, then submit it to see the LB score.

In the initial step of both methods, we performed 5-fold cross-validation using accuracy as the evaluation metric for classification. The following sub-sections will present the cross-validation results for both methods across each region.

3.1 Afghanistan

The tree-based models achieved the following accuracies: 0.862 for LGBM, 0.853 for CatBoost, and 0.857 for the weighted average ensemble models. In contrast, the InceptionTime model achieved a lower accuracy of 0.614.

Figure 5 and figure 6 showed the top 15 features based on their importance for LGBM and CatBoost models. It's worth noting that the most important feature is VH polarization band from Sentinel-1 data in February. Moreover, several related SAR bands and features appear within the top 15, especially for the CatBoost classifier. This suggests that Sentinel-1 data is more powerful for crop mapping in winter due to its ability to go through clouds.

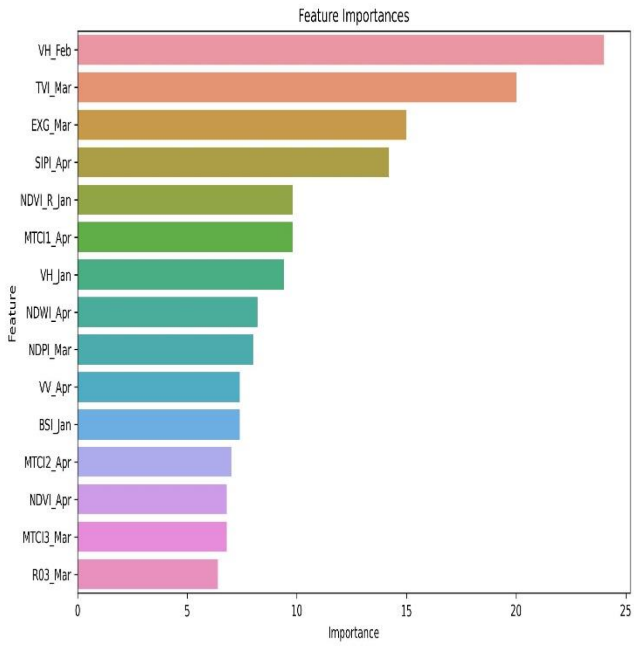

Fig5. – Feature importance for LGBM classifer over Afghanistan.

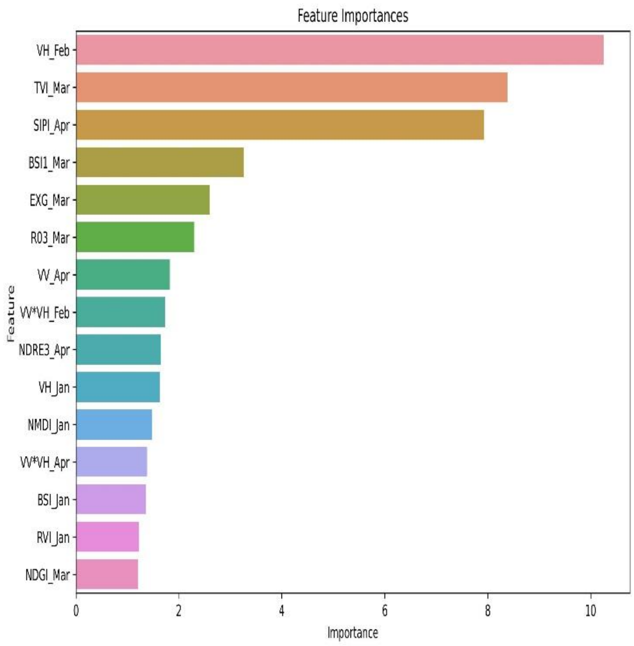

Fig6. – Feature importance for CatBoost classifer over Afghanistan.

3.2 Sudan

The tree-based models achieved high accuracies: 0.976 for LGBM, 0.966 for CatBoost, and 0.971 for the weighted average ensemble models. In contrast, the InceptionTime model achieved a lower accuracy with 0.614. Figures 7 and 8 demonstrate that the top four predictors are related to Sentinel-2 data, which were extracted during the summer months, such as July when cloud cover is less frequent, thus minimizing its impact on the performance of optical data. And in general, we have more important features from sentinel-2 due to the long timespan.

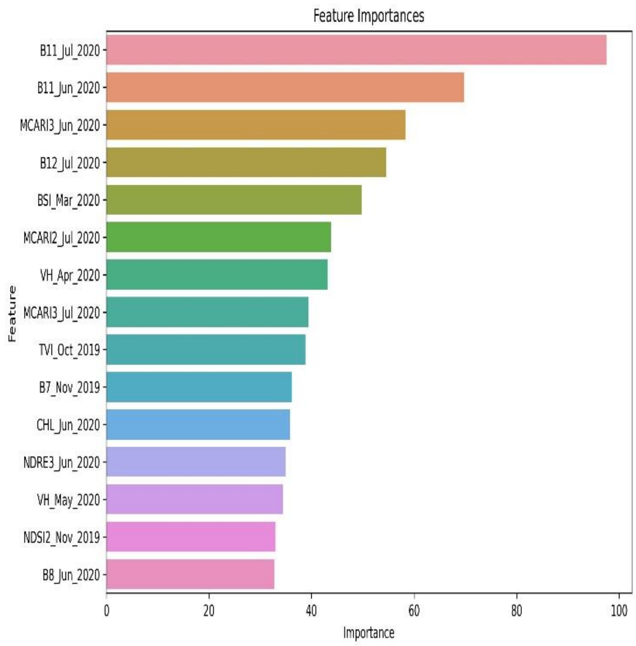

Fig7. – Feature importance for LGBM classifer over Sudan.

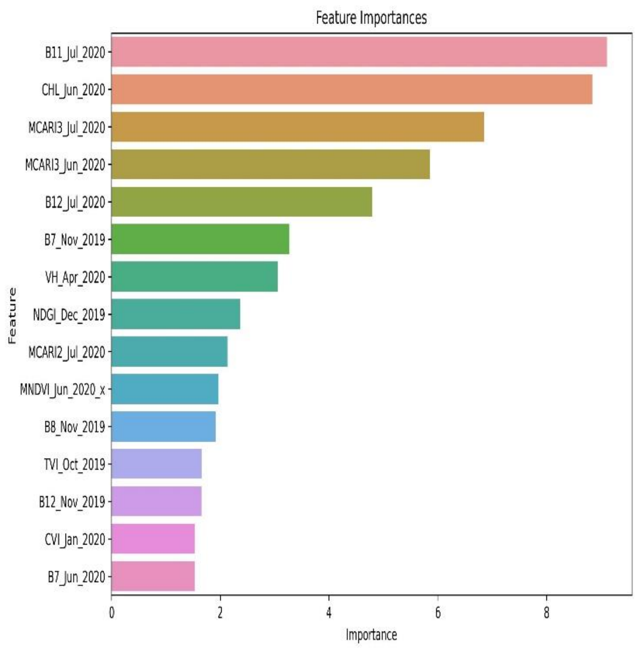

Fig8. – Feature importance for CatBoost classifer over Sudan.

3.3 Iran

The tree-based models achieved notable accuracies: 0.970 for LGBM, 0.952 for CatBoost, and 0.961 for the weighted average ensemble models. In contrast, the InceptionTime model achieved a lower accuracy by 0.598.

Figures 9 and 10 demonstrate that soil moisture features, such as leaf area index and land evapotranspiration flux from the SMAP Level-4 collection, strongly influence cropland predictions over Iran.

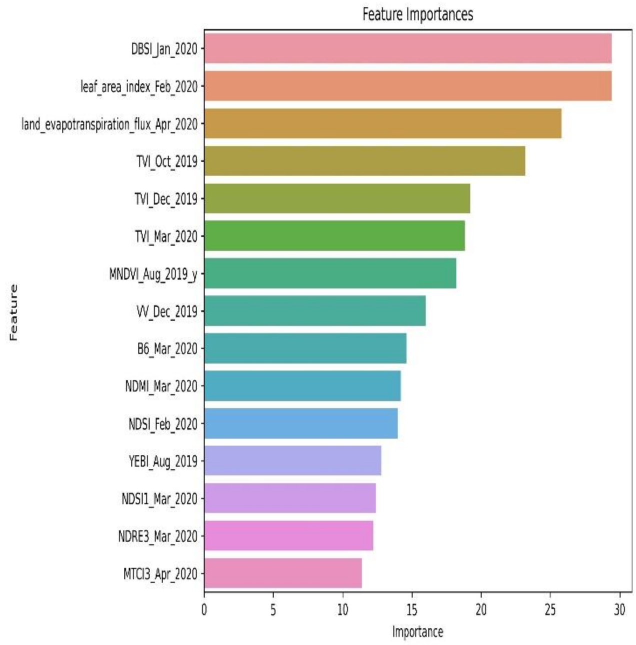

Fig9. – Feature importance for LGBM classifer over Iran.

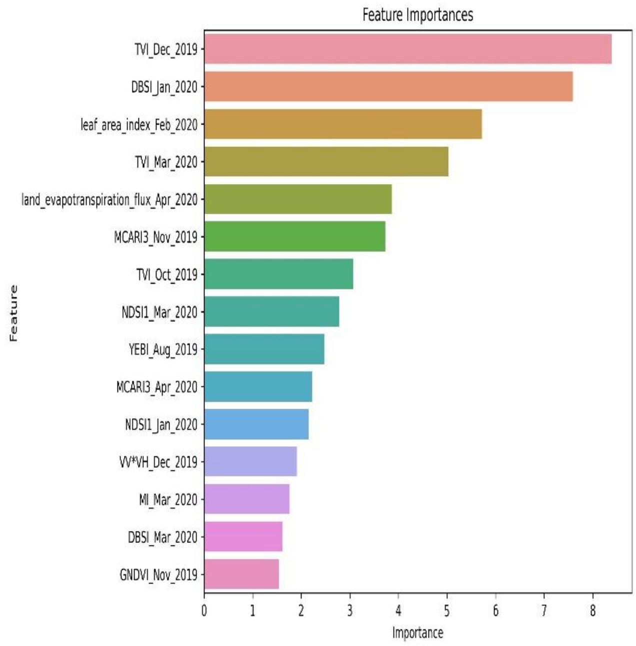

Fig10. – Feature importance for CatBoost classifer over Iran.

In the final step, we concatenated the predictions from individual regions into one file and evaluated the accuracy on the Leaderboard (LB). The weighted ensemble, using LGBM and CatBoost, outperformed the InceptionTime model with an accuracy of 0.94, while the InceptionTime model achieved an accuracy of 0.857.

In conclusion, weighted ensemble model using tree-based models (LGBM and CatBoost) and tabular data structure outperformed timeseries model (InceptionTime model) using 3-d data. Also, Long timeseries span improved the cropland maps accuracy, as explained above 1-year time interval achived 0.961 and 0.971 accuracy for Iran and Sudan, respectively. While 4-months interval achieved 0.857 accuracy.

4. CONCLUSION

In conclusion, the weighted ensemble model, incorporating tree-based models (LGBM and CatBoost) with tabular data, outperformed the time-series model (InceptionTime) using 3D data. Additionally, the use of longer time series intervals significantly improved cropland mapping accuracy. Specifically, the 1-year time interval achieved accuracies of 0.961 and 0.971 for Iran and Sudan, respectively, while the 4-month interval resulted in an accuracy of 0.857. Additionally, the importance of Sentinel-1 data for cropland mapping, particularly during the winter months, is evident due to its ability to penetrate cloud cover.

REFERENCES

[1] Yu, Q., Shi, Y., Tang, H., Yang, P., Xie, A., Liu, B., & Wu, W. (2017). eFarm: a tool for better observing agricultural land systems. Sensors, 17(3), 453.

[2] Matton, N., Sepulcre Canto, G., Waldner, F., Valero, S., Morin, D., Inglada, J., ... & Defourny, P. (2015). An automated method for annual cropland mapping along the season for various globally-distributed agrosystems using high spatial and temporal resolution time series. Remote Sensing, 7(10), 13208-13232.

[3] Tatsumi, K., Yamashiki, Y., Torres, M. A. C., & Taipe, C. L. R. (2015). Crop classification of upland fields using Random forest of time-series Landsat 7 ETM+ data. Computers and Electronics in Agriculture, 115, 171-179.

[4] Kussul, N., Lavreniuk, M., Skakun, S., & Shelestov, A. (2017). Deep learning classification of land cover and crop types using remote sensing data. IEEE Geoscience and Remote Sensing Letters, 14(5), 778-782.

[5] Zhang, M., Wu, B., Zeng, H., He, G., Liu, C., Tao, S., ... & Liu, Y. (2021). GCI30: a global dataset of 30 m cropping intensity using multisource remote sensing imagery. Earth System Science Data, 13(10), 4799-4817.

[6] Mazzia, V., Khaliq, A., & Chiaberge, M. (2019). Improvement in land cover and crop classification based on temporal features learning from Sentinel-

2 data using recurrent-convolutional neural network (R-CNN). Applied Sciences, 10(1), 238.

[7] Shelestov, A., Lavreniuk, M., Vasiliev, V., Shumilo, L., Kolotii, A., Yailymov, B., ... & Yailymova, H. (2019). Cloud approach to automated crop classification using Sentinel-1 imagery. IEEE Transactions on Big Data, 6(3), 572-582.

[8] Orlíková, L., & Horák, J. (2019, May). Land cover classification using sentinel-1 SAR data. In 2019 international conference on military technologies (ICMT) (pp. 1-5). IEEE.

[9] Brinkhoff, J., Vardanega, J., & Robson, A. J. (2019). Land cover classification of nine perennial crops using sentinel-1 and-2 data. Remote Sensing, 12(1), 96.

AUTHORS

Mohammad Alasawedah

I am an Earth observation and climate change researcher, currently working within a research group focused on advanced computing for earth observation at Eurac research. My work involves using advanced computing techniques to collect and analyze data on environmental factors.