environmental sciences proceedings

MDPI

Proceeding paper

Utilizing Machine Learning and Time Series Sentinel-1&2 Imagery to Map Rice Fields for Sustainable Agriculture Using the Google Earth Engine System in Mazandaran Province, Iran †

Mahdiyeh Fathi, Reza Shah-Hosseini *, and Sara Salahi

School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran; mahdiyeh.fathi@ut.ac.ir (M.F.); saraasalahi@gmail.com (S.S.)

* Correspondence: rshahosseini@ut.ac.ir

† Presented at the 5th International Electronic Conference on Remote Sensing, 7–21 November 2023; Available online: https://ecrs2023.sciforum.net.

Abstract: The spatial distribution of rice fields in the Mazandaran province is most importance in understanding various crucial aspects such as food security, water usage, greenhouse gas emissions, and disease transmission. However, the lack of an accessible ground truth map in the region poses a challenge. To overcome this, we employed a pixel-based paddy rice mapping (PPPM) algorithm to generate an accurate ground truth map by identifying flooding signals. In this study, we proposed a novel method using the Random Forest (RF) classification and time-series imagery from Sentinel-1 and Sentinel-2 within the Google Earth Engine system to effectively differentiate rice fields from other crop lands. The proposed method involved several steps. Firstly, we extracted various bands (such as blue, green, red, red edge 1/2/3, NIR, n-NIR) and essential indices (including NDVI, LSWI, DVI, RVI, WDRVI, SAVI, EVI, VARIGREEN, and GNDVI) from Sentinel-2 imagery. Next, we performed multi-collinearity analysis to select the optimal indices and bands, which included NIR, Red Edge 3, NDVI, LSWI, VARI-green, DVI, and GNDVI. In the subsequent step, monthly composites of the optimal indices and bands were generated from March to August. The RF classification algorithm was then applied to classify the study area into six classes: water, crop land, urban, forest, outcrop, and range land. Finally, the Radar Vegetation Index was extracted from Sentinel-1 imagery to accurately separate the rice fields from the other crop lands. Our approach had impressive results, with a high accuracy and kappa coefficient of 99% and 98%, respectively, for rice fields. This information is crucial for policymakers and researchers, as it enables them to make informed decisions regarding food security, water usage, greenhouse gas emissions, and disease transmission in the Mazandaran province. Additionally, it provides valuable insights into the expansion of crop lands through agricultural irrigation. By utilizing machine learning techniques and satellite imagery, we can generate accurate and reliable information about cropland areas, facilitating the development of effective strategies for sustainable agriculture and food security.

Keywords: Rice Fields; Machine Learning; Sentinel-1 and Sentinel-2; Google Earth Engine

Citation: To be added by editorial staff during production.

Academic Editor: Firstname Lastname

Published: date

Copyright: © 2023 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Rice is the main food for over half of the world's population. The growth of rice fields depends on water and heat. The area and spatial distribution of rice fields are critical in terms of water usage, food security, and greenhouse gas emissions. Given the scarcity of water resources, understanding the monitoring of rice fields is essential [1,2].In recent years, remote sensing technologies such as optical and radar imagery (Sentinel-1/2, Landsat, MODIS, etc.) have played an important role in identifying rice fields as they expand [3].

The extracted time series maps from optical satellite imagery (Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), Soil Adjusted Vegetation Index (SAVI) and Normalized Difference Water Index (NDWI), etc.) can be used to differentiate rice fields from other crops. The extracted phenology from these maps plays a crucial role in monitoring crop growth stages [1]. By analyzing the changes in vegetation indices or backscattering values over a specific period, we can determine phenological parameters such as the beginning and end of the growing season and the highest value of the vegetation index. Rice fields are flooded before planting, and the vegetation index has low values. However, the vegetation index increases with crop growth. Therefore, the extracted indices from optical images aid in estimating phenological parameters [4].

Various methods have been introduced to identify rice fields using satellite imagery, including Phenology-Based methods, Machine-Learning methods, and Deep-Learning methods [2]. For example, Fathi et al, proposed using Fusion of In-Decoder CNN and Data Augmentation Techniques to identify rice fields using extracted maps of Land Surface Temperature (LST), Normalized Difference Vegetation Index (NDVI), and Land Surface Water Index (LSWI) from multi-temporal Landsat-8 imagery [2]. Peng et al, proposed using C-band multi-temporal Radarst-2 imagery to extract quad-polarized backscattering coefficients, Cloude-Pottier and Freeman-Durden decomposition parameters as random forest (RF) classification features to identify rice fields [5]. Fiorillo et al, integrated Sentinel-1/2 imagery to extract Normalized Difference Vegetation index (NDVI) and backscattering values to identify rice fields using the Random Forest (RF) algorithm [6]. Waldini et al, employed a Phenology-Based approach using Sentinel-2 imagery to extract Normalized Difference Vegetation Index (NDVI) and Normalized Difference Water Index (NDWI) to identify rice fields [7]. Fernando et al, proposed a Google Earth Engine (GEE) based multi-index random forest (RF) classification approach to map rice fields using Landsat imagery [8]. Zhai et al, used multi-sensor remote sensing data (Radarsat-2 and Sentinel-2 imagery) and the Random Forest algorithm. The results of this research showed that the red-edge band and red-edge vegetation index are better than the NIR (near-infrared) band and NDVI (Normalized Difference Vegetation Index) to map rice fields [9]. Liu et al, developed an algorithm based on feature optimization and random forest modeling, using 35 common remote sensing features to construct 7 feature combinations using Sentinel-2 imagery [10].

Identifying rice crops, like other agricultural crops, presents several challenges. One of the most common challenges is cloud cover in rice planting areas when using optical imagery. Additionally, it can be difficult to identify rice fields with an area smaller than the spatial resolution of the sensor used. Another challenge is identifying rice fields in sloping areas when using radar imagery due to the terrain's shape. Moreover, the accuracy of identifying rice fields from radar imagery is generally lower compared to optical imagery. Separating rice fields from rainfall crops can also be a challenge. Additionally, the spectral similarity of rice with other classes, rice with different species, and mixed pixels can make identification complex. Also, spectral bands are affected by radiometric distortions [11]. To overcome these challenges, a possible solution is to combine radar imagery with optical images that have high resolution. This approach can help improve the accuracy of identifying rice fields [2,3].

In this study, our objective is to investigate the effectiveness of combining Sentinel-1 and Sentinel-2 imagery to differentiate rice fields from other crop fields using the Random Forest classifier within the Google Earth Engine system in Mazandaran province, Iran. We extracted several features from Sentinel-2 imagery and utilized the Multi-Collinearity matrix to identify the optimal features. Finally, we produced a 6-class map using the Random Forest classifier. Additionally, Radar Vegetation index (RVI) was extracted from Sentinel-1 imagery to distinguish rice fields from other types of crop Land. Finally, our findings demonstrates the potential of combining Sentinel-1 and Sentinel-2 imagery to accurately identify rice fields in Mazandaran province, Iran.

2. Materials and Methods

2.1. Study Area

The study was located in Mazandaran province, Iran, in 2022. This province possesses the necessary weather conditions and accessible water sources for rice cultivation. The Ministry of Agriculture provides the agricultural calendar to the public annually. According to the rice planting calendar specific to Mazandaran province, rice cultivation begins in March, and the crop is harvested in mid-August. Figure 1 illustrates the study area.

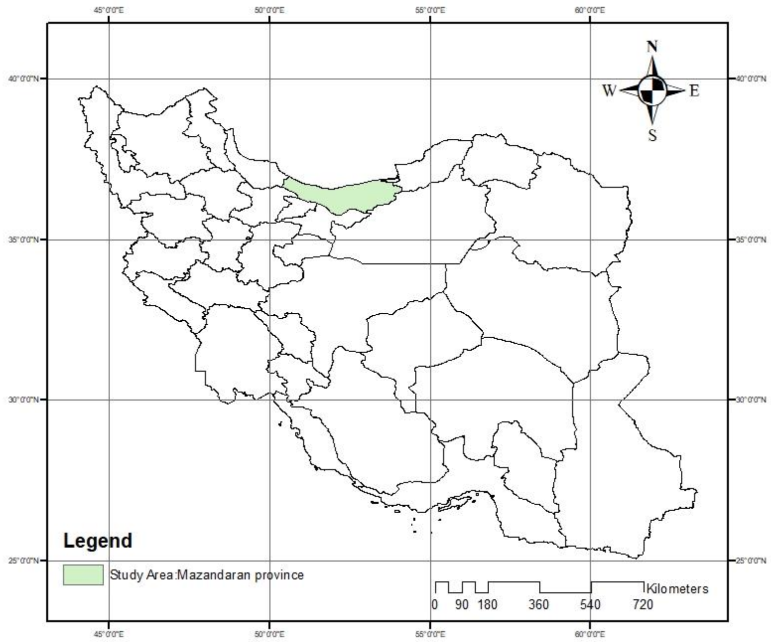

Figure 1. Study Area.

2.2. Dataset

This study used a variety of dataset to map rice fields, including Sentinel-1 SAR, and Sentinel-2. Due to the unavailability of a ground truth map in Mazandaran province, we utilized the proposed phenology-based approach by Dong et al [1]. This approach was corrected based on the rice cultivation calendar in Mazandaran province and was used to generate a ground truth map.

2.3. Methodology

We propose a novel method that utilizes Random Forest (RF) Classifier and time series Sentinel-1&2 imagery within the Google Earth Engine system to distinguish rice fields from other types of crop lands. Figure 2 depicts our proposed method. In the subsequent sections, we will provide a detailed description of our proposed approach.

2.3.1. Feature Selection

Sentinel-2 images were utilized to extract various spectral indices, such as NDVI, LSWI, DVI, RVI, WDRVI, SAVI, EVI, VARIGREEN, and GNDVI. These indices, along with spectral bands including blue, green, red, red edge 1/2/3, NIR, and n-NIR, were used as input for a Correlation-Based Multi-Collinearity Matrix to select the most suitable features [12]. Optimal features were chosen based on having less than 10% correlation. Finally, monthly mosaics of optimal features were generated as input for the random forest classifier.

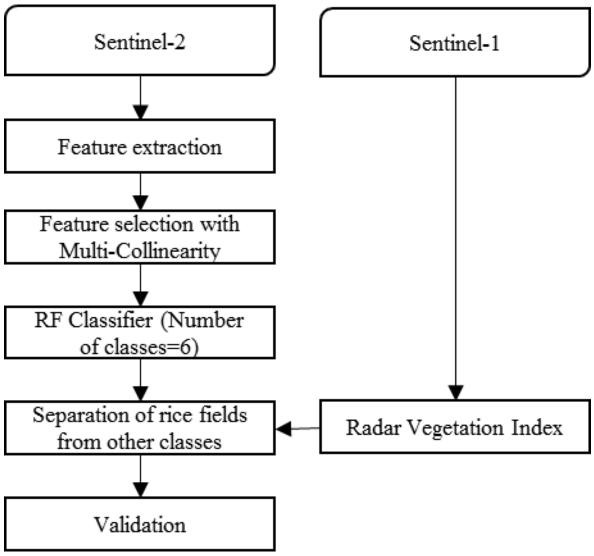

Figure 2. Flowchart of the proposed method.

2.3.2. Random Forest Classifier

Random Forest is a machine learning algorithm extensively employed for classification and regression. It belongs to the ensemble learning category, which combines multiple decision trees to generate predictions. In a Random Forest, a group of decision trees is created, with each tree constructed using a random subset of the training data and a random subset of the available features. During the training process, each tree independently produces predictions, and the final prediction is determined by aggregating the results from all the trees [13]. We utilized the Random Forest algorithm to classify the study area into six distinct classes: water, crop land, urban, forest, outcrop, and range land. In the next step, we separated the rice fields from the remaining crop land by employing the extracted Radar Vegetation Index (RVI) from Sentinel-1 images.

2.3.3. Generate Rice Field Map Using Radar Vegetation Index (RVI)

After generating the classified map of the study area with six classes, we utilized the extracted Radar Vegetation Index (RVI) (Equation 1) from Sentinel-1 to distinguish rice fields from the crop land class. Initially, we generated monthly RVI mosaics. Subsequently, we calculated the successive differences of these monthly mosaics and composited them together to effectively separate rice fields from other crop land areas.

$$\mathrm{RVI} = \frac{4\sigma_{\mathrm{VH}}}{\sigma_{\mathrm{VV}} + \sigma_{\mathrm{VH}}} \tag{1}$$

Where σ_{VH} and σ_{VV} represent the measured backscattering intensities [14].

3. Results

To produce a classified map of our study area, we encountered the obstacle of not having a reliable ground truth map. To overcome this challenge, we utilized satellite imagery and relied on the interpretation of multiple samples to train the Random Forest (RF) classifier.

Our first step was to establish sampling points for both training and testing purposes. These points were then utilized to train the RF algorithm. Figure 3 depicts the study area and the location of the sampling points for each class. We allocated 80% of the sampling points for training and 20% for test.

Next, we extracted monthly mosaics of 9 spectral features and 8 spectral bands from Sentinel-2 images. To ensure optimal feature selection, we applied a Multi-Collinearity Matrix based on correlation. This resulted in the identification of seven features: NIR, Red Edge 3, NDVI, LSWI, VARI-green, DVI, and GNDVI.

Figure 3. Study area and sampling points.

After the algorithm was trained using optimal features, and training samples, a classified map consisting of six classes - water, crop land, urban, forest, outcrop, and range land - was produced for the Mazandaran province. The classified map was validated by utilizing a confusion matrix and a sampling test. Table 1 shows the results of the confusion-matrix. Figure 4 depicts classification map using RF algorithm.

Table 1. The results of the confusion-matrix.

Class	Urban	Water	Forest	Out crop	Range land	crop land
Recall	75%	66%	100%	80%	100%	83%
Precision	60%	100%	100%	88%	81%	91%
Overall Accuracy				87%		
Kappa Coefficient				84%		

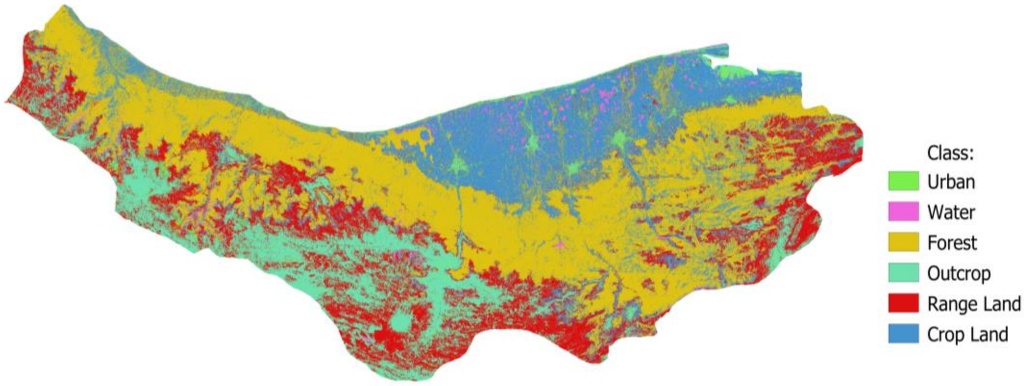

Figure 4. Generated classification map using RF algorithm.

After classifying the study area into 6 classes, we proceeded to separate the rice fields from the remaining crop land. This was achieved by utilizing a generated composite from monthly Radar Vegetation Index mosaics. Our approach arrived impressive results with a high accuracy and kappa coefficient of 99% and 98%, respectively, for rice fields. Figure 5 depicts the distribution of rice fields on the map.

(**a**)

(**b**)

Figure 5. (**a**) Map of generated rice using proposed method, and (**b**) Ground truth map.

4. Conclusion

To overcome the challenge of insufficient ground truth data in Mazandaran province, a solution was proposed utilizing the Google Earth Engine system. We utilized Sentinel-2 images and the Random Forest (RF) algorithm to classify the Mazandaran province into six distinct classes. The classified map has reached 87% overall accuracy and 84% kappa coefficient. Subsequently, Sentinel-1 images were employed to specifically identify rice fields from other crop land in classified map. Our approach arrived satisfactory overall accuracy (99%) and kappa coefficient (98%) values, indicating that the extracted Radar Vegetation Index (RVI) from Sentinel-1 images was effective in successfully separating rice planting areas from other crop land.

Author Contributions: Conceptualization, M.F., R.SH.; and S.S.; Methodology, M.F., R.SH., and S.S.; Project administration, R.SH.; Resources, M.F., R.SH., and S.S.; Validation, M.F., R.SH., and S.S.; Supervision, R.SH.; Writing—original draft, M.F., R.SH., and S.S.; Writing-review and editing, M.F., R.SH., and S.S.

Funding: This article was not received funding.

Acknowledgments: The authors would like to thank the Ministry of Agriculture Jihad, Iran for providing the agricultural calendar to the public annually.

References

1. Dong, J.; Xiao, X.; Menarguez, M.A.; Zhang, G.; Qin, Y.; Thau, D.; Biradar, C.; Moore III, B. Mapping paddy rice planting area in northeastern Asia with Landsat 8 images, phenology-based algorithm and Google Earth Engine. *Remote sensing of environment* **2016**, *185*, 142-154. [CrossRef]
2. Fathi, M.; Shah-Hosseini, R. Automatic rice fields mapping by fusion of in-decoder CNN and data augmentation techniques on Landsat-8 multi-temporal images. *Earth Observation and Geomatics Engineering* **2021**, *5*, 74-95. [CrossRef]
3. Mosleh, M.K.; Hassan, Q.K.; Chowdhury, E.H. Application of remote sensors in mapping rice area and forecasting its production: A review. *Sensors* **2015**, *15*, 769-791. [CrossRef]
4. Wang, J.; Huang, J.-f.; Wang, X.-z.; Jin, M.-t.; Zhou, Z.; Guo, Q.-y.; Zhao, Z.-w.; Huang, W.-j.; Zhang, Y.; Song, X.-d. Estimation of rice phenology date using integrated HJ-1 CCD and Landsat-8 OLI vegetation indices time-series images. *Journal of Zhejiang University. Science. B* **2015**, *16*, 832. [CrossRef]
5. Peng, W.; Li, S.; He, Z.; Ning, S.; Liu, Y.; Su, Z. Random forest classification of rice planting area using multi-temporal polarimetric Radarsat-2 data. In Proceedings of the IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium, 2019; pp. 2411-2414. [CrossRef]
6. Fiorillo, E.; Di Giuseppe, E.; Fontanelli, G.; Maselli, F. Lowland rice mapping in Sédhiou Region (Senegal) using sentinel 1 and sentinel 2 data and random forest. *Remote Sensing* **2020**, *12*, 3403. [CrossRef]
7. Waldini, H.; Shidiq, I.P.A.; Rokhmatuloh, R.; Supriatna, S. Rice Crop Phenology Model to Monitor Rice Planting and Harvesting Time using Remote Sensing Approach. In Proceedings of the E3S Web of Conferences, 2021; p. 03020. [CrossRef]
8. Fernando, W.A.M.; Senanayake, I. Developing a two-decadal time-record of rice field maps using Landsat-derived multi-index image collections with a random forest classifier: A Google Earth Engine based approach. *Information Processing in Agriculture* **2023**. [CrossRef]
9. Zhai, P.; Li, S.; He, Z.; Deng, Y.; Hu, Y. Collaborative mapping rice planting areas using multisource remote sensing data. In Proceedings of the 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, 2021; pp. 5969-5972. [CrossRef]

10. Liu, Y.; Xiao, D.; Yang, W. An algorithm for early rice area mapping from satellite remote sensing data in southwestern Guangdong in China based on feature optimization and random Forest. *Ecological Informatics* **2022**, *72*, 101853. [CrossRef]
11. Moghimi, A.; Celik, T.; Mohammadzadeh, A. Tensor-based keypoint detection and switching regression model for relative radiometric normalization of bitemporal multispectral images. *International Journal of Remote Sensing* **2022**, *43*, 3927-3956. [CrossRef]
12. Alin, A. Multicollinearity. *Wiley interdisciplinary reviews: computational statistics* **2010**, *2*, 370-374. [CrossRef]
13. Pal, M. Random forest classifier for remote sensing classification. *International journal of remote sensing* **2005**, *26*, 217-222. [CrossRef]
14. Dey, S. Radar vegetation index code for dual polarimetric Sentinel-1 data in EO browser. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.