



OPEN

## Evaluating the efficiency of coarser to finer resolution multispectral satellites in mapping paddy rice fields using GEE implementation

Mirza Waleed<sup>1</sup>, Muhammad Mubeen<sup>1</sup>✉, Ashfaq Ahmad<sup>2</sup>, Muhammad Habib-ur-Rahman<sup>3,11</sup>✉, Asad Amin<sup>4</sup>, Hafiz Umar Farid<sup>5</sup>, Sajjad Hussain<sup>1</sup>, Mazhar Ali<sup>1</sup>, Saeed Ahmad Qaisrani<sup>1</sup>, Wajid Nasim<sup>6</sup>, Hafiz Muhammad Rashad Javeed<sup>1</sup>, Nasir Masood<sup>1</sup>, Tariq Aziz<sup>7</sup>, Fatma Mansour<sup>8</sup> & Ayman EL Sabagh<sup>9,10</sup>✉

Timely and accurate estimation of rice-growing areas and forecasting of production can provide crucial information for governments, planners, and decision-makers in formulating policies. While there exists studies focusing on paddy rice mapping, only few have compared multi-scale datasets performance in rice classification. Furthermore, rice mapping of large geographical areas with sufficient accuracy for planning purposes has been a challenge in Pakistan, but recent advancements in Google Earth Engine make it possible to analyze spatial and temporal variations within these areas. The study was carried out over southern Punjab (Pakistan)-a region with 380,400 hectares devoted to rice production in year 2020. Previous studies support the individual capabilities of Sentinel-2, Landsat-8, and Moderate Resolution Imaging Spectroradiometer (MODIS) for paddy rice classification. However, to our knowledge, no study has compared the efficiencies of these three datasets in rice crop classification. Thus, this study primarily focuses on comparing these satellites' data by estimating their potential in rice crop classification using accuracy assessment methods and area estimation. The overall accuracies were found to be 96% for Sentinel-2, 91.7% for Landsat-8, and 82.6% for MODIS. The F1-Scores for derived rice class were 83.8%, 75.5%, and 65.5% for Sentinel-2, Landsat-8, and MODIS, respectively. The rice estimated area corresponded relatively well with the crop statistics report provided by the Department of Agriculture, Punjab, with a mean percentage difference of less than 20% for Sentinel-2 and MODIS and 33% for Landsat-8. The outcomes of this study highlight three points; (a) Rice mapping accuracy improves with increase in spatial resolution, (b) Sentinel-2 efficiently differentiated individual farm level paddy fields while Landsat-8 was not able to do so, and lastly (c) Increase in rice cultivated area was observed using satellite images compared to the government provided statistics.

Rice, grown on 12% of the global cropland area<sup>1</sup>, is a major staple food for half of the world's population. Climate change impacts such as global temperatures and changing precipitation dynamics and land cover changes due to urbanization and industrialization negatively affect paddy rice (*Oryza sativa L.*) production. Rice, an important food crop of Pakistan, is also an export product making a valuable foreign exchange. In Pakistan US\$1,376

<sup>1</sup>Department of Environmental Sciences, COMSATS University Islamabad, Vehari Campus, Islamabad 61100, Pakistan. <sup>2</sup>Asian Disaster Preparedness Center (ADPC), Bangkok, Thailand. <sup>3</sup>Institute of Crop Science and Resource Conservation (INRES), Crop Science, University of Bonn, 53115 Bonn, Germany. <sup>4</sup>Queensland Alliance for Agriculture and Food Innovation, Hertely Tackle Building (Bldg 83), Glasshouse Road, St. Lucia, QLD 4072, Australia. <sup>5</sup>Department of Agricultural Engineering, Bahauddin Zakariya University, Multan, Pakistan. <sup>6</sup>Department of Agronomy, Faculty of Agriculture and Environment Sciences, The Islamia University of Bahawalpur (IUB), Punjab, Pakistan. <sup>7</sup>Department of Earth and Environmental Sciences, University of Waterloo, Waterloo, ON N2L 3G1, Canada. <sup>8</sup>Business and Economics Faculty, Department of Economics, Siirt University, Siirt, Turkey. <sup>9</sup>Department of Agronomy, Faculty of Agriculture, Kafrelsheikh University, Kafrelsheikh, Egypt. <sup>10</sup>Department of Field Crops, Faculty of Agriculture, Siirt University, Siirt, Turkey. <sup>11</sup>Department of Agronomy, Muhammad Nawaz Shareef-University of Agriculture, Multan, Pakistan. ✉email: muhammadmubeen@cuivehari.edu.pk; mhabibur@uni-bonn.de; habibagri@hotmail.com; aymanelsabagh@gmail.com

million were earned by exporting rice alone in 2015–16<sup>2</sup>. However, on the downside, rice is a rigorous water crop when associated with other cereal crops. Therefore, water productivity counts a lot for a developing country like Pakistan, which is already a water-stressed country<sup>3</sup>. Consequently, Small-scale farmers, intensify cropping cycles and increase water use to enhance crop production. Paddy crop, with its high water consumption, can also be linked to global water security issues in the future<sup>4</sup>. At the same time paddy fields are a source of methane (CH<sub>4</sub>), and CH<sub>4</sub> emissions from paddy rice cultivation have increased from 7 to 8% worldwide between years 2000 and 2016<sup>5</sup>. Since paddy rice fields provide habitat for free-ranging ducks and wild waterfowls, they contribute to avian influenza virus<sup>1</sup>. Therefore, a regular, accurate and appropriate mapping of paddy crop is crucial for achieving food security, tackling climate change, controlling disease transmission, and improving crop management and production<sup>6</sup>. Remote sensing techniques are extensively used to map rice-growing areas at a frequent temporal resolution and to analyze the effects of external variables (e.g., water shortage and climate change) on rice crop<sup>7–9</sup>. Remote sensing provides valuable information about agricultural progress and seasonal crop variation, directly linked to the life cycle of paddy rice<sup>10</sup>. Instead of field investigation, which is time consuming, labor intensive, and financially expensive, remote sensing with standardized approaches can be effectively used to analyze the spatial and temporal variability that may affect the life cycle and rice crop health<sup>11</sup>.

Southern Punjab (Pakistan), a region vulnerable to climate change, has elevated poverty levels<sup>12</sup>, and lacks resources to meet its people's needs<sup>13</sup>. Inadequate crop management practices and advances in climatic variability are main threats to agricultural production in this province. Therefore the region is expected to face elevated poverty levels in the near future. Rice crop, which is the second major crop of Kharif season in terms of area after cotton, is the second most consumed staple food in southern Punjab after wheat<sup>14</sup>. For effective management of rice production and demand, accurate satellite instruments need to be identified so that effective mapping techniques can be implemented in the region<sup>15</sup>. Government officials, agronomists, and researchers often require accurate and timely crop information for producing statistical reports, policy implementation, and advancement in precision agriculture<sup>15</sup>.

Google Earth Engine (GEE), a cloud based repository of satellite imagery and processing capabilities has gained popularity in recent years due to its ability to handle large datasets, and its power to analyze and visualize big data. This platform enables a user to compute and analyze worldwide satellite imagery with its complete spatial and temporal characteristics in an internet-based cloud platform provided free of cost by Google<sup>16</sup>. GEE even strongly supports studies related to earth and its processes. With its onset in 2010, GEE delivers users all satellite imagery, cloud-based computing, and machine learning algorithms to easily process large datasets<sup>17</sup>. In previous studies<sup>1,6,18,19</sup>, researchers have benefited from GEE platform for paddy rice mapping using phenology-based algorithms along with time series of vegetation indices.

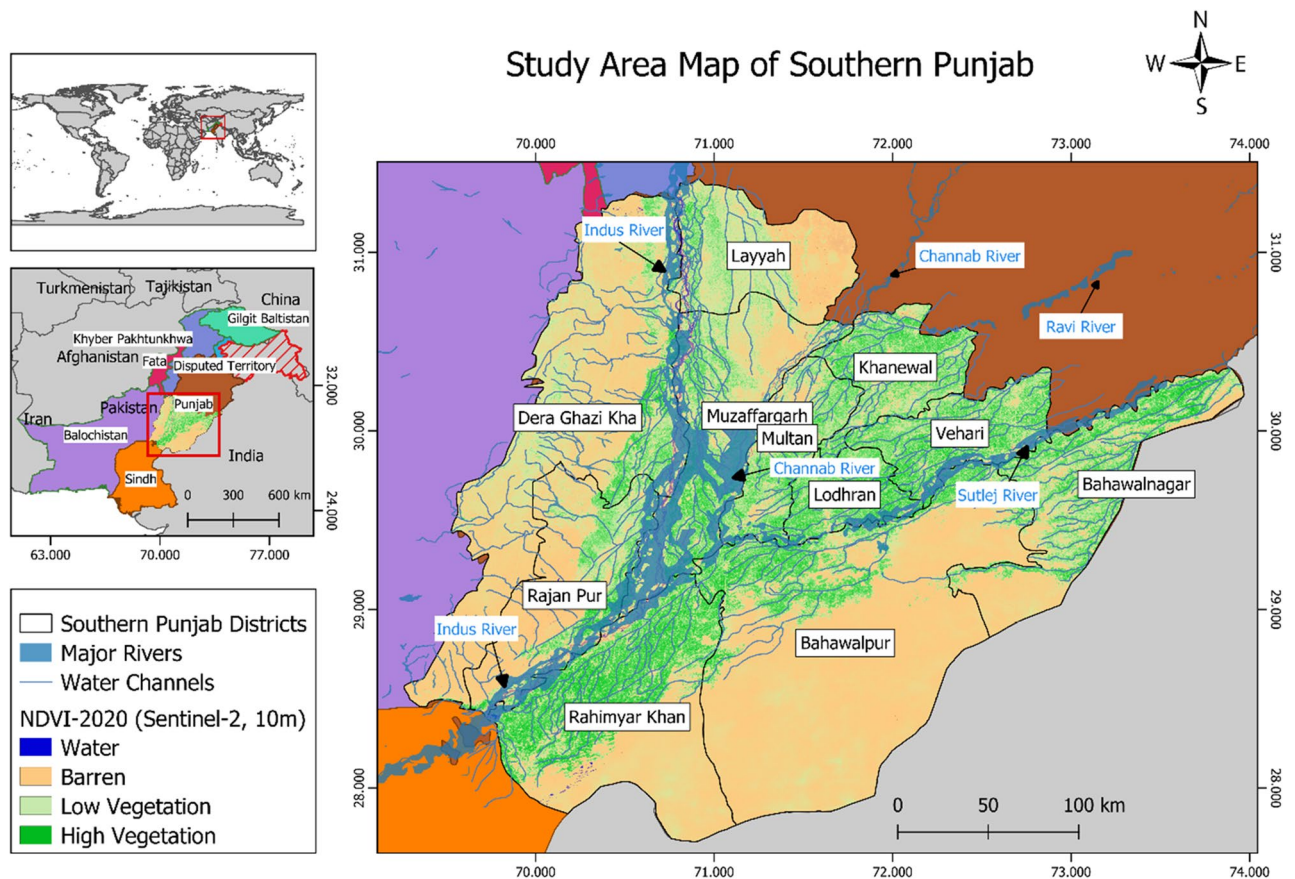
Earlier studies<sup>20–23</sup> support the individual capability of Sentinel-2, Landsat-8 and MODIS multispectral instruments for rice crop classification. However, to date, no study has compared the efficiency of these multispectral instruments in classifying individual paddy rice fields at farm level. Therefore, this study mainly focuses on using the GEE cloud platform to classify rice crops in southern Punjab region (comprised of Multan, Dera Ghazi Khan, and Bahawalpur divisions). The objective is to compare the freely available and GEE integrated coarser to finer resolution multispectral instruments for rice classification using the RF classifier. In this, classification results will be compared based on their accuracy assessment and area estimation. The results will highlight the comparative efficiency of freely available satellite datasets for mapping paddy crop in the region. Furthermore, the resulting area estimations will play a significant role in precision agriculture, eventually helping the decision-making processes.

## Material and methods

**Study area.** The study area is southern Punjab region (in Pakistan) that lies between latitude 27.30° N to 31.50° N and longitude 69.20° E to 73.55° E (Fig. 1). It has three administrative divisions: Bahawalpur, Multan, and Dera Ghazi Khan. Bahawalpur division has three, while Multan and Dera Ghazi Khan divisions have four districts each<sup>12</sup>. Four major rivers, Chenab, Indus, Sutlej, and Ravi, flow through the study area. Based on visual interpretation, the NDVI derived map shows that the Multan division has higher vegetation than the other two. While Bahawalpur, Dera Ghazi Khan, Rajan Pur, and Layyah districts showed sparse vegetation trends. The southern Punjab region lacks sufficient resources to support its people as the majority of areas are undeveloped with weak infrastructure<sup>13</sup>. The gravity of the situation can be understood by the fact that out of 40 million people in Pakistan living under poverty line, with 10 million of those belong to the Southern Punjab region<sup>12</sup>. In the region, most of the population depends on agriculture for their livelihoods. Few districts, such as Bahawalnagar and Vehari, contain a dense network of water channels, making them suitable for rice production. Additionally, the Bahawalnagar district had high water table than other districts, which makes it fit for rice crop in the Kharif and not suitable for other crops such as cotton<sup>24</sup>. In southern Punjab, there are two main cropping seasons: Kharif (April–June, Summer) and Rabi (October–December, Winter). Rice is the dominant crop of the Kharif season<sup>4</sup>.

According to the crop report for the year 2020 provided by the Crop Reporting Service, Department of Punjab (<http://www.crs.agripunjab.gov.pk/reports>), the total rice area estimation of the Multan, D. G. Khan, and the Bahawalpur division was 380,400 hectares (Fig. 2). The statistical data show that the Bahawalnagar district had the highest area estimate of 87,810 hectares. The lowest area of 13,760 hectares of rice was reported in Rajan Pur and Layyah districts. We use these estimates for validation of area estimation in our study.

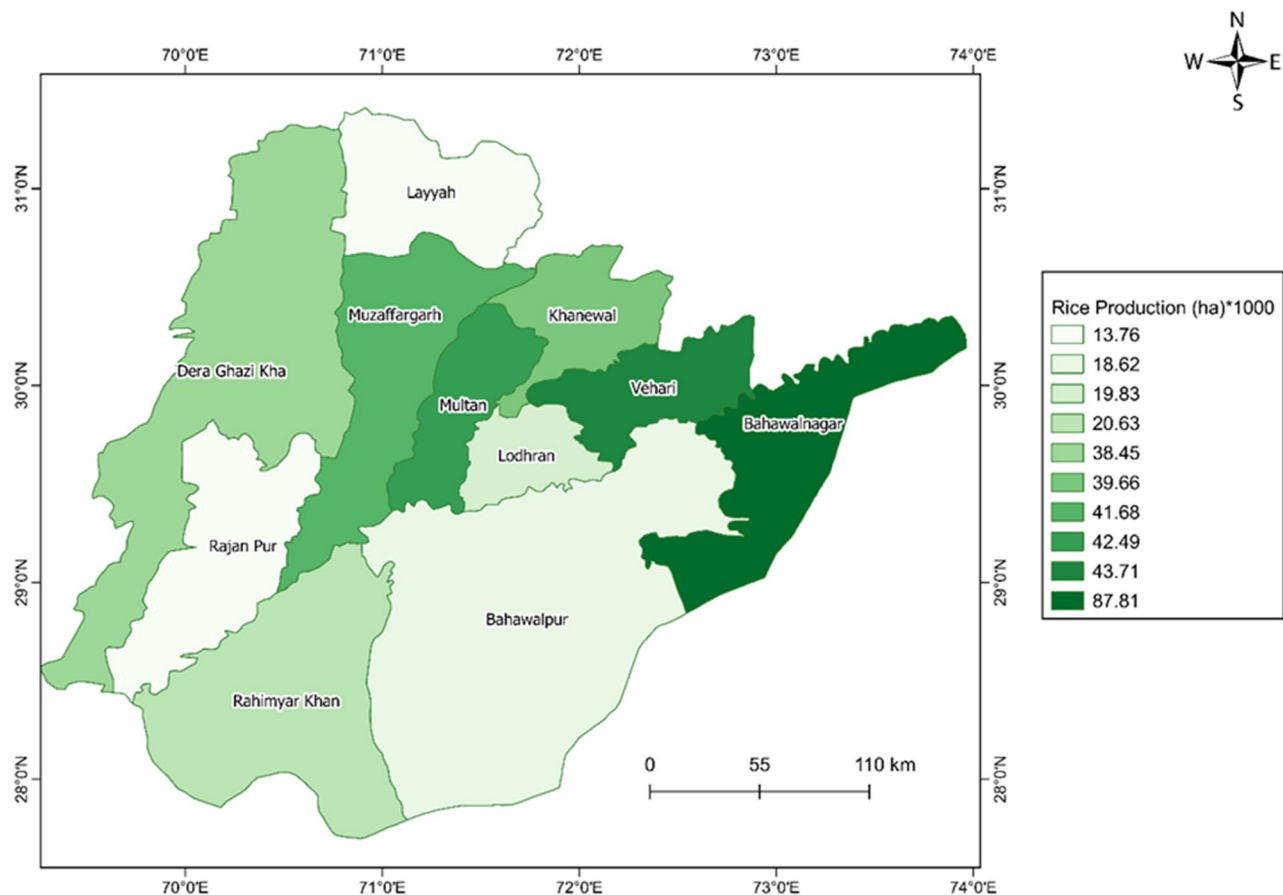
**Data.** Satellite imagery in the Near Infrared (NIR), Shortwave-Infrared (SWIR) and microwave region provide valuable information about soil moisture, improving the identification of paddy rice fields<sup>20</sup>. Singha et al.<sup>6</sup> and Wang et al.<sup>20</sup> studies conclude that a combination of normalized difference vegetation index (NDVI) and



**Figure 1.** Map of the study area (Southern Punjab). The map is designed in ArcGIS Pro V2.9 software available at ESRI website (<https://www.esri.com/en-us/arcgis/products/arcgis-pro/overview>). The boundary shapefile used to draw these maps are available at Humdata website (<https://data.humdata.org/dataset/cod-ab-pak>).

Sentinel 1 backscatter coefficient (measure of reflective strength of radar body) along with Land Surface Temperature (LST) of day and night time and precipitation data can be effectively use to identify the rice fields during the flooding/transplanting and ripening phases with an overall accuracy greater than 80 percent. Data derived from Sentinel-1 provides valuable information in the microwave region and is used to differentiate paddy rice fields in cloud prone regions because of its capability to penetrate cloud cover. On the other hand, areas without cloud cover issue rely on NIR and SWIR data from multispectral instruments. Similarly, GEE platform is delivering users with coarse to moderate resolution imagery, such as Moderate Resolution Imaging Spectroradiometer (MODIS) at 250 m resolution, Landsat-8 at 30 m resolution, and Sentinel-2 at 10 m resolution. These aforementioned satellite instruments are used in previous studies for paddy rice mapping<sup>1,20–23,25–27</sup>. Sentinel-2 Satellites (2nd series of satellites after Sentinel-1) are part of the Copernicus program initiated by the European Space Agency (ESA)<sup>7</sup>. This satellite mission carries multispectral scanners and consists of two satellites which are Sentinel-2A and Sentinel-2B<sup>28</sup>. In Sentinel-2 satellites, multispectral imaging instruments (MSI) are installed with the ability to record 13 wide-swath bands<sup>29</sup>. Depending upon the band, each satellite provides 10–60 m of spatial resolution with a temporal resolution of 5 days using both satellite constellations. NASA's Landsat-8 mission is a new generation satellite that carries the legacy of Landsat missions. This satellite gets spectral information in visible (V), NIR, SWIR, and Thermal Infrared (TIR) regions<sup>1</sup>. In this satellite mission, two sensors are installed: Operational Land Imager (OLI) and TIR. The OLI obtains spectral information in nine different spectral bands, whereas TIR records thermal information in two bands<sup>30,31</sup>. MODIS mission contains two satellites: Terra and Aqua. These two satellites have nearly the same sensors. The mission monitors the Earth surface every one to two days and obtains spatial data in 36 different spectral bands between 0.405 and 14.385  $\mu\text{m}$ . This mission provides 250 m to 1 km spatial resolution coverage in visible and infrared regions with a 2330 m wide swath<sup>32</sup>.

For preparing paddy rice seasonal spectral profile, Sentinel-2 derived NDVI at 10 m resolution<sup>33</sup> and Sentinel-1 Backscatter coefficient time-series variation<sup>6</sup> were accessed and four observations per month (depending upon satellite revisit intervals) were obtained from 10 randomly selected rice fields inside the Southern Punjab boundary. These training points were then digitized inside GEE, and then their monthly average was calculated individually. For average monthly temperature, MODIS 250 m resolution LST dataset was used<sup>34</sup>, and separate day and night LST were calculated. For precipitation, Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) pentad (5-day) precipitation dataset<sup>35</sup> was used to calculate average monthly precipitation. The average of LST and precipitation were calculated as average values for the southern Punjab inside GEE.

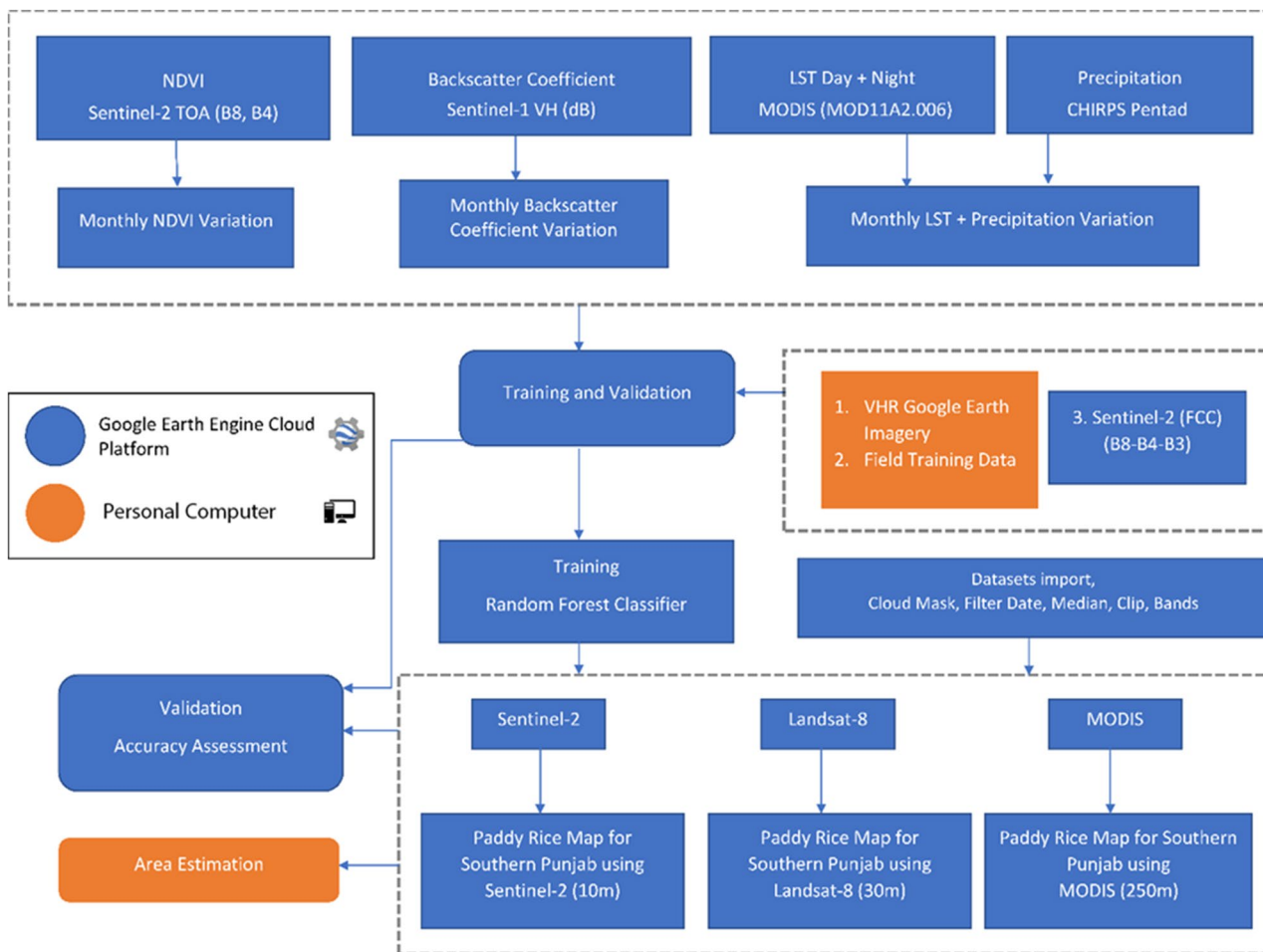


**Figure 2.** Ground truth data for rice production in southern Punjab for year 2020 (Reported by Crop Reported Service, Government of Punjab). The map is designed in ArcGIS Pro V2.9 software available at ESRI website (<https://www.esri.com/en-us/arcgis/products/arcgis-pro/overview>). The data used in this map is available at CRS website (<http://www.crs.agripunjab.gov.pk>).

**Methods.** The methodology is divided into three parts. In the first part, rice growth profiles are evaluated using multiple instruments (Sentinel-1, Sentinel-2, Landsat-8, and MODIS), which are crucial to crosscheck training points (non-rice, rice, and water). In second part of methodology, training points (rice, and non-rice) are used to calibrate RF classifier and classify the land cover. After classification, the last part compares derived paddy rice maps using accuracy assessments and differences in the estimated actual areas. The overall methodology performed is shown in Fig. 3. Moreover, all the steps performed in this methodology (i.e., pre-processing and classification) are according to the guidelines and regulations provided by the United States Geological Survey (USGS) (<https://www.usgs.gov>), and GEE (<https://www.explorer.earthengine.google.com/terms>).

**Random Forest (RF) based Classification.** Random Forest (RF), a machine learning classifier, has widely been used in several studies for paddy rice mapping<sup>36–40</sup>. Because of its high precision, RF (having a usage percentage of 49% globally in satellite image classification) is the most popular classifier, along with Classification and Regression Tree (13%) and Support Vector Machine (11%)<sup>41</sup>. The RF classifier contains a mixture of tree classifiers in which every individual classifier is created using an arbitrary vector tested independently from the user given input, and each of the trees in the classifier gives a vote for the most excellent prevalent class to classify the given input vector<sup>20</sup>. One of the main advantages of an RF classifier is that it needs two types of parameters to be set, while in contrast, the Support Vector Machine need more than two input parameters by the user. Also, the RF classifier can deal with the data without values, which is not conceivable with Support Vector Machine. Another essential advantage of this classifier is that it also offers the comparative importance of different features during the classification process, which is very valuable in selecting features<sup>42</sup>.

Hyperparameter optimization was performed initially which chooses the optimal parameters to be used for algorithm to yield highest accuracy in a given situation. As a result of hyperparameter optimization, max number of trees value was set to 115 which yield maximum accuracy for our study area. Then the trained classifier was used further to classify paddy rice pixels. For classification validation, the training classifier was fed with validation data, and then the resulting trained classifier was used to calculate the error matrix for resultant maps. All these steps, including cloud masking, data filtering, median, clipping, mosaicking, classifier training, classification, and classification validation, were performed individually for each satellite dataset inside GEE.



**Figure 3.** Methodology chart for rice crop classification.

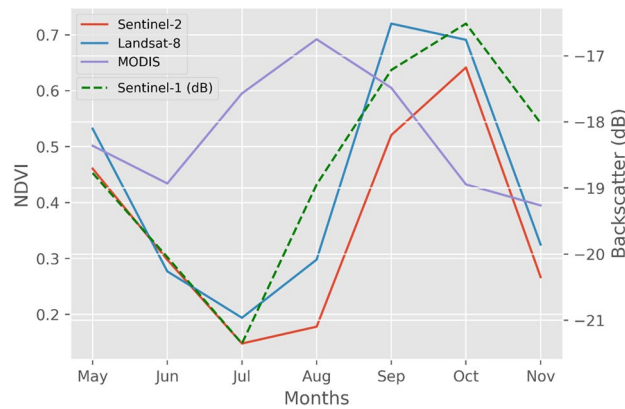
Flooding or transplanting phase		High vegetation period			Harvesting period	
May	June	July	August	September	October	November

**Table 1.** Rice crop calendar in southern Punjab, Pakistan.

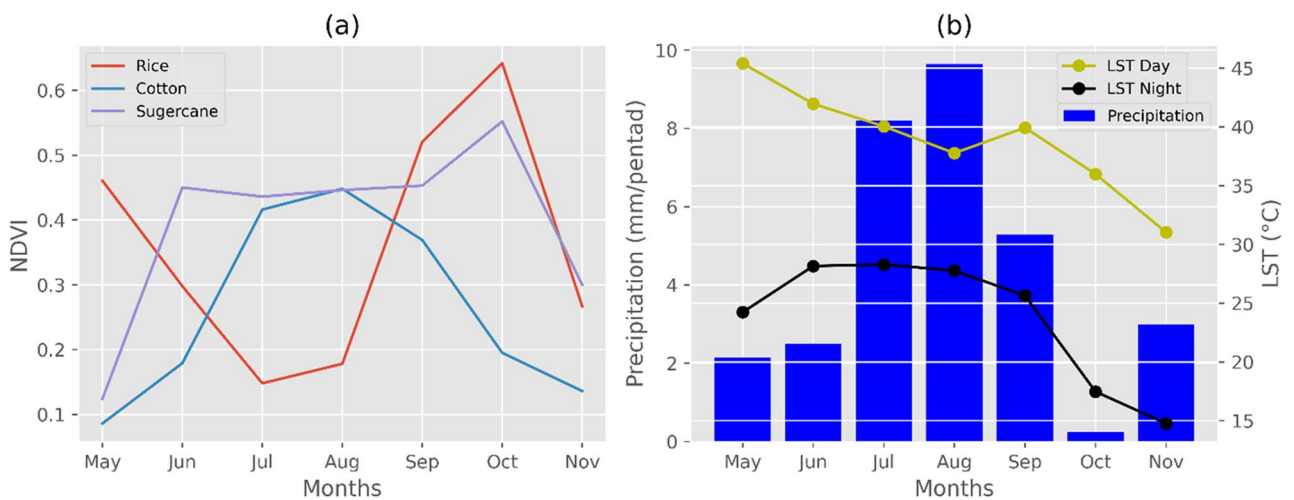
**Temporal variation of rice using NDVI and backscatter coefficient.** Since values of NDVI and Sentinel-1 backscatter coefficient vary with the variation in rice growth stages, they were used inside GEE to track rice growth stages for 2020<sup>6</sup>. Rice crop goes through three distinct growth stages: (1) flooding or transplanting phase; (2) high vegetation period; (3) the harvest period<sup>4,14,43,44</sup>. In the Punjab region, rice growing starts from the mid-May with harvesting from September to November<sup>4,45</sup>. The phenology of rice is shown in Table 1. These development stages are the major reason for variation in NDVI and Sentinel-1 backscatter values.

To map the temporal variation of rice, we used a literature based methodology<sup>6</sup>. The monthly time series of NDVI was created using Sentinel-2 (10 m), Landsat-8 (30 m), and MODIS (250 m) satellite data (Fig. 4). For comparison, the backscatter coefficient (dB) time series was created using Sentinel-1 in dual-band cross-polarization with vertical transmit, horizontal receive (VH) (Fig. 4). The NDVI variation was observed for rice using three multispectral instruments (Fig. 4). Given the rice crop sowing and transplanting phase, NDVI values of Sentinel-2 and Landsat-8 are low (~0.16 to 0.20) in months of June and July. However, the values are comparatively higher (~0.51 to 0.55) in September and October, representing higher vegetation. For MODIS, peak values of NDVI were observed in August. From Fig. 4, it was observed that backscatter values were low (~ -18 to -21 dB) for the flooding stage of rice due to the presence of water and for the harvest stage due to surface scatter from the open land. In contrast, they were significantly high (~ -16.51 dB) for the vegetation stage due to volume scattering in the rice plant.

NDVI variation of various Kharif season crops (Rice, Cotton, and Sugercane) using Sentinel-2 MSI data were compared (Fig. 5a). From their variation, it was observed that NDVI values of crops differ spatially and temporally, therefore, provide great potential in differentiating the crop's growth stages. This phenomenon was



**Figure 4.** Average NDVI temporal profile for rice using Sentinel-2, Landsat-8 and MODIS derived NDVI data and average Sentinel-1 backscatter coefficient (dB) temporal profile for rice.

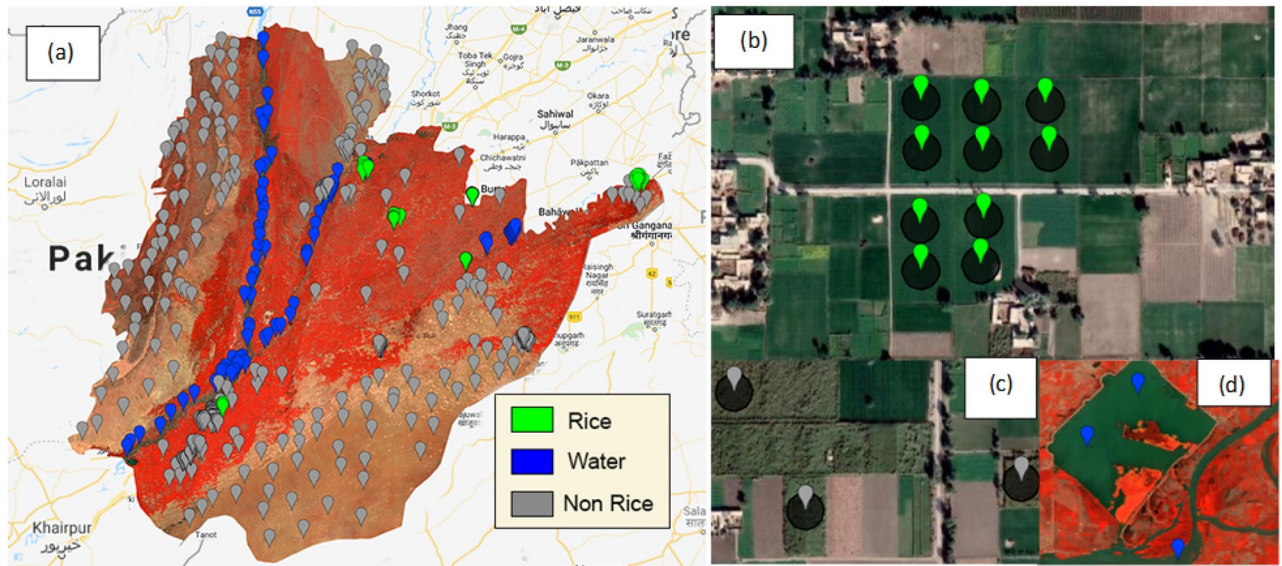


**Figure 5.** Crops NDVI temporal variation using Sentinel-2 (a) and Average Monthly LST (Day and Night) along with CHIRPS precipitation temporal variation (b), for 2020 Kharif season in Southern Punjab region.

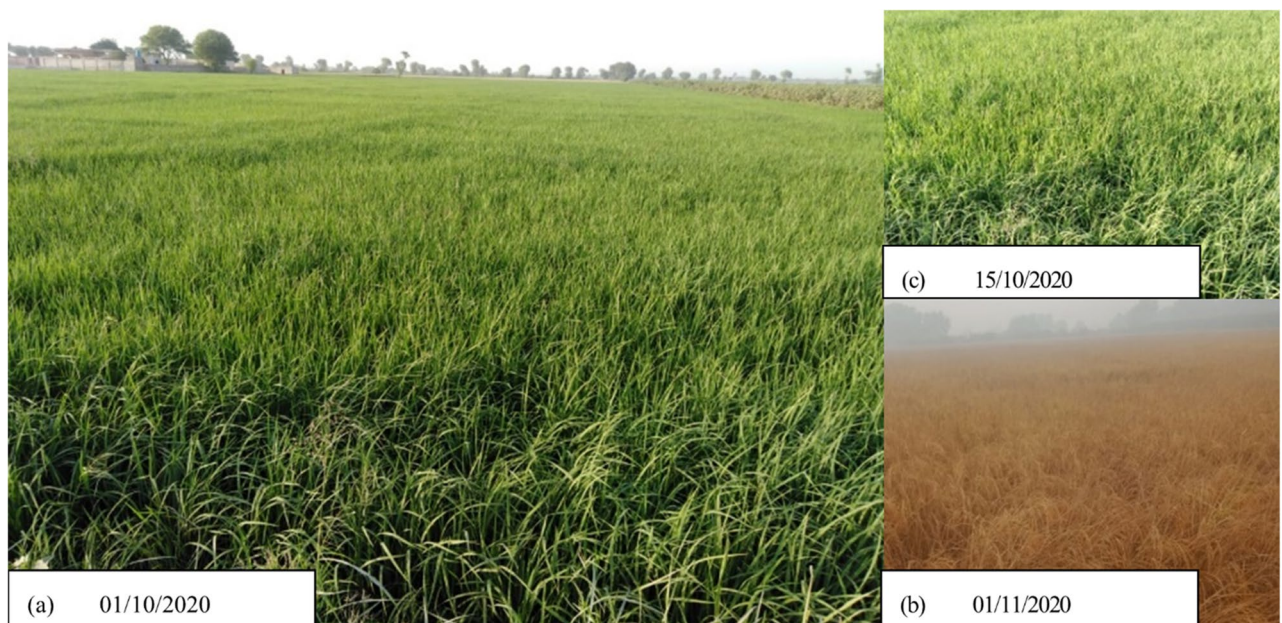
also confirmed by previous studies<sup>46–50</sup>. In the last, Fig. 5b shows average day and night LST derived from MODIS satellite along with CHIRPS precipitation data. We used Landsat-8 (100 m thermal data) to derive LST for both day and night. Precipitation time series was prepared using CHIRPS data. CHIRPS is a precipitation dataset which provides 35 years of grided precipitation data with a spatial resolution of 5.5 km<sup>51</sup>. The purpose of LST and precipitation data (Fig. 5b) was solely to verify the climatic condition for the year 2020 which helped us in planning survey and collecting field samples, in appropriate growth cycle of rice crop in souther Punjab. Figure 5b shows that mean precipitation was high during July to September with moderate mean LST. Literature shows that such weather condition is favorable for enhanced rice growth<sup>52</sup>.

**Training samples.** Sentinel-1 derived Synthetic Apertur Radar (SAR) data was used to check training points before feeding them to classifier for classification. Previous studies have reported, that the SAR backscatter value is sensitive to water<sup>21</sup>. As paddy rice follows different growth stages, SAR based backscatter can be used to track any pixel either paddy or non paddy. Same principle was applied in this study to differentiate non-rice training points from paddy rice. For specifically non-rice class, each training point was cross checked using Sentinel-1 SAR backscatter data. Only samples which does not follow spectral trend as shown by paddy rice (Fig. 4) were used for representing non-rice training class. To ensure proper coverage and correct selection of the ground truth data for our large study area, we collected the ground truth data (using random sampling approach) through multiple sources: (1) Very High Resolution VHR images from the Google Earth, (2) 167 field rice latitude longitude coordinates point samples randomly collected during the rice field survey in October–November, and (3) Visual interpretation of Sentinel-2 false color composite (FCC) at 10 m spatial resolution.

Firstly, we used the Sentinel-2 FCC and VHR Google Earth images to digitize points by visual interpretation except for rice (Fig. 6a). Secondly for each class, we made areas of interest AOIs as circle buffers of the points with the radius of 20 m for Sentinel-2 and Landsat-8 (Fig. 6b–d). Thirdly, for rice, we took accurate field coordinates



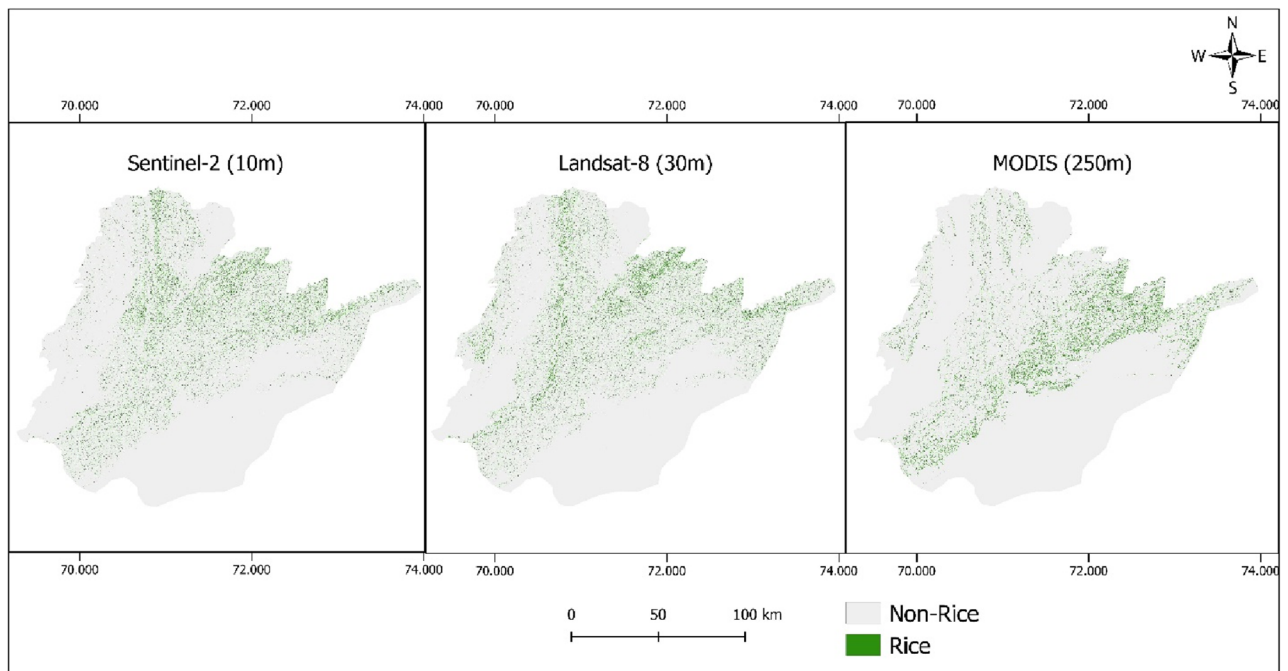
**Figure 6.** Training samples in the study area, (a) all samples, (b) ricefield samples, (c) non-rice samples, and (d) water samples. The map is designed in Google Earth Engine, which is an cloud computing browser based platform (<https://earthengine.google.com>).



**Figure 7.** Ricefield photographs, taken during training sample collection at different locations and timeperiod within the study area. The photographs (a–c) are taken by first author, during the field survey.

from 6 locations of Southern Punjab and photographs of rice fields at different time intervals are shown in Fig. 7. For collecting rice coordinate samples, we choose 6 random locations in the study area which are shown in Fig. 6a (note the 6 areas where rice points are clustered) and collected samples using random sampling approach. For MODIS, however, it was observed that due to its coarser spatial resolution of 250 m, 20 m buffers was not suitable for training points. So, for MODIS, training points without buffer were used and each pixel contributing to paddy rice class was digitized based on its centroid.

The AOIs without clear land cover information were omitted. The VHR Google Earth imagery and false-color composite combination for Sentinel-2 at 10 m and field photographs are used to clarify the land cover types, i.e., water, barren, forests, non-paddy, and urban areas. The classes, including non-paddy agriculture, forests, grasslands, urban and barren soil, were merged in one class named non-rice. Water class was kept separate due to its distinct spectral characteristic and was merged with non-rice class after training classifier. The reason for separating water class initially from other land-use classes was to restrict water pixels to create confusion during



**Figure 8.** Rice crop map in southern Punjab for the year 2020 using Sentinel-2, Landsat-8 and MODIS datasets.

classification<sup>53</sup>. In this study, separating water pixels from paddy rice was crucial as they both share similar spectral characteristics<sup>6</sup> (only when rice is in flooding growth stage). The field sizes, shapes, and proximities were considered in labeling land cover information of AOIs.

A total of 690 AOIs were collected, including 167 rice samples, 88 water samples, and 435 non-rice samples from different locations of Southern Punjab. The AOIs were distributed randomly, covering the entire study area. The generated AOIs were randomly split (by category) into a 70/30 ratio; 70% were used as training data, while the remaining 30% were used for validation. The paddy rice classification was performed individually for each multispectral satellite dataset utilizing the same training AOIs and the same year (2020).

**Pre-processing.** Sentinel-2, Landsat-8, and MODIS datasets (freely available inside the GEE data catalog) were used separately and individually for paddy rice classification. Parameters that were kept the same for a fair comparison among these three datasets were (1) same training, and validation data, (2) same satellite data acquisition dates were used (July–October), and the median was taken to ensure best pixel coverage<sup>6</sup>, (iii) Similar spectral bands (if present) were utilized, (4) The same RF classifier was used in each classification, and in the last, (5) same procedure was applied for accuracy assessment and area estimation.

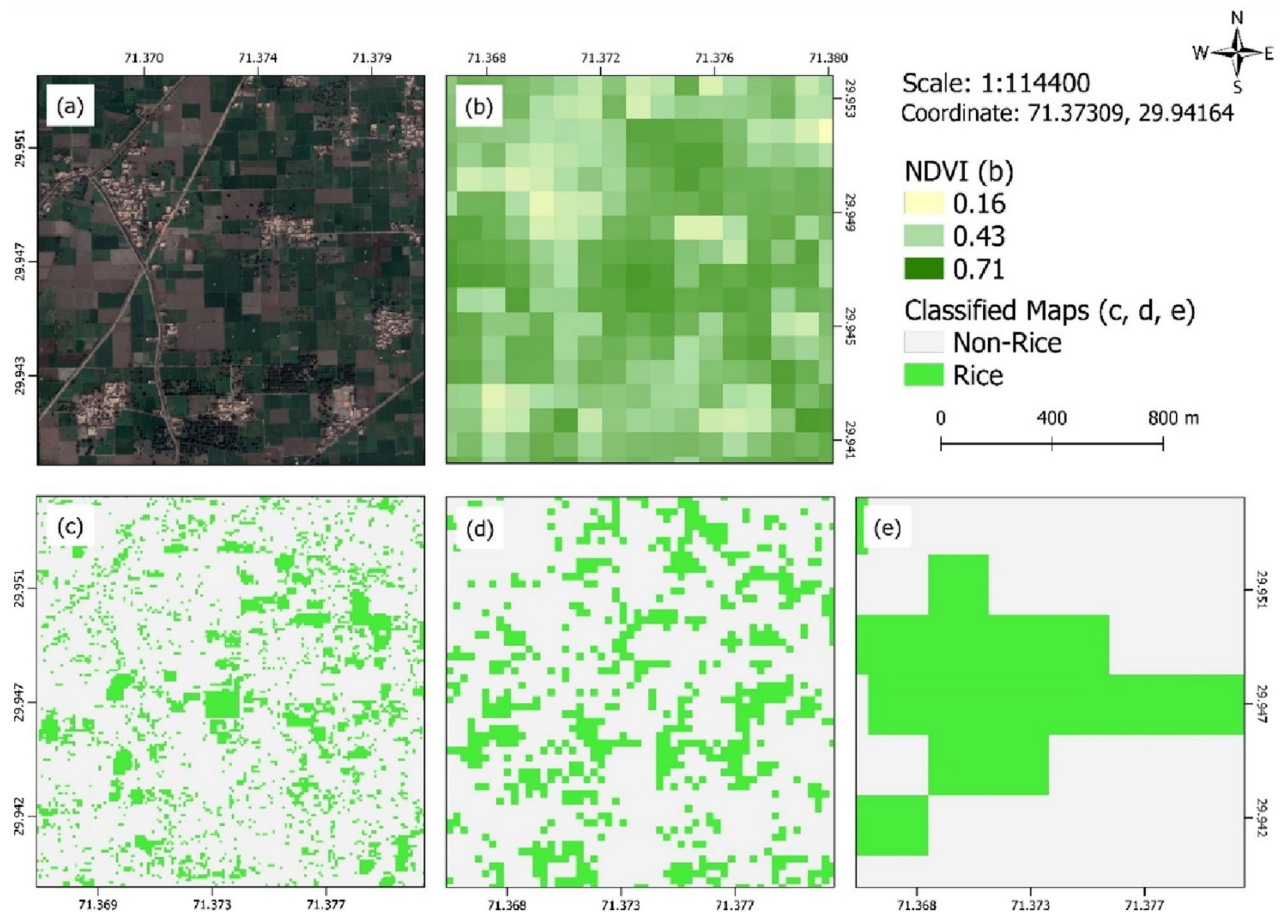
These datasets were used inside GEE. First, these datasets were used with a cloud mask individually, ensuring minimum cloud coverage in data<sup>54</sup>. Then images between July and October were filtered, and their median was taken. Then the datasets were clipped to the boundaries of the study area. For Sentinel-2 and Landsat-8, Visible band with NIR and SWIR bands are used, because NIR and SWIR bands of Sentinel-2 and Landsat-8 enhance crop classification<sup>55</sup>. However, for MODIS, band-1 SWIR and band-2 NIR images with a spatial resolution of 250 m are used<sup>56</sup>.

## Results

**Rice maps using Sentinel-2, Landsat-8 and MODIS.** The rice classified maps for Southern Punjab are generated using Sentinel-2, Landsat-8, and MODIS satellite data (Fig. 8). From their visual comparison, it is observed that MODIS derived maps form clusters instead of classifying individual rice fields, which can be attributed to a coarse-resolution of 250 m. As a result, most pixels of the MODIS map are misclassified because it mixed neighboring crops, and subsequently formed large clusters indicating them rice. Another problem observed in MODIS classified map is denser rice classified pixels clustered in the northeast of southern Punjab, especially in Bahawalnagar, Bahawalpur, and Rahim Yar Khan districts (Fig. 8). This is because of dense vegetation in those areas that can be visually observed in Fig. 1. When compared with Fig. 1, it is confirmed (through visual interpretation) that those areas were actually other crops (Fig. 1) which MODIS has classified totally as rice (Fig. 8).

Landsat-8, on the other hand, performed better than MODIS. At some points, it was successful in classifying individual rice fields. Paddy crop is mostly grown in small, segmented fields; therefore, spatial resolution plays a great role in identifying individual paddy fields<sup>57</sup>. Landsat-8 misclassified some pixels as rice clusters in Khanewal and Dera Gazi Khan districts, that can be clearly seen in the Fig. 8. The rest of the Landsat-8 data is reliable enough for further use. Sentinel-2 classified individual paddy rice with greatest precision as compared to the other two datasets, and successfully differentiating individual paddy rice fields at a resolution of 10 m. The





**Figure 9.** Comparing classified maps at a random location in Southern Punjab using VHR Google Earth imagery (a), Sentinel-2 10 m derived NDVI (b), rice classified map using Sentinel-2 (c), Landsat-8 (d) and MODIS (e).

Satellites	Class	Error matrix		UA (%)	PA (%)	F1-score (%)	Kappa (%)	OA (%)
		Non-rice	Rice					
Sentinel-2	Non-rice	1712	39	97.7	97.7	97.7	81.2	96.0
	Rice	40	205	84	83.7	83.8		
Landsat-8	Non-rice	1683	26	97.8	98.5	98.1	73.3	91.7
	Rice	38	189	77.9	73.3	75.5		
MODIS	Non-rice	1663	28	93.5	98.3	95.8	61.2	82.6
	Rice	116	137	73	54.1	65.5		

**Table 2.** Error matrix of paddy rice classification for three satellite data (UA user's accuracy, PA producer's accuracy, OA overall accuracy).

performance of these datasets in differentiating paddy rice fields is presented in Fig. 9. For visual comparison, VHR Google Earth imagery (Fig. 9a) and Sentinel-2 derived NDVI at 10 m resolution (Fig. 9b) were used as reference images while derived rice classified maps for Sentinel-2 (Fig. 9c), Landsat-8 (Fig. 9d), and MODIS (Fig. 9e) were compared. At a scale of 1:114,400, a random point (with coordinates 71.37309, 29.94164) was selected and the corresponding derived maps (Fig. 9c–e) were compared at this view. It shows that MODIS formed a large cluster and is not able to identify individual paddy fields. On the other hand, Landsat-8 has identified some individual paddy fields, however, it has misclassified the majority of neighboring small pixels as rice. Comparatively, Sentinel-2 has efficiently classified individual rice fields with lowest number of misclassified pixels. Because of its high resolution of 10 m, Sentinel-2 classified individual paddy fields better than both the Landsat-8 and MODIS.

**Technical validation.** Through visual inspection (Figs. 8, 9) and accuracy assessment (Table 2), we find that the MODIS-derived map is not reliable for differentiating individual paddy fields and cannot be used in

local scale regions decision-making. As for the nearly similar area of Sentinel-2 and MODIS datasets, Sentinel-2 derived map was justified through accuracy assessment and visual interpretation. However, MODIS failed to withstand accuracy assessment and visual interpretation mainly due to cluster formation and huge misclassification.

Accuracy assessment is important to check the reliability of land use classification and error matrix is a common method to assess the accuracy<sup>58</sup>. An error matrix is also used to draw other statistical measures of accuracy such as overall accuracy (OA), commission error (i.e., user's accuracy (UA)), omission error (i.e., producer's accuracy (PA)), and the kappa coefficient (K)<sup>59</sup>. The UA evaluated by dividing the number of correctly classified pixels (each category) by the total number of classified pixels in that land use category (the row total)<sup>60</sup>. The UA actually shows the probability that the particular pixel of land-use class of classified map actually represent that land-use class on ground. The equation used to derive UA is given as Eq. 1.

$$UA = \frac{\text{Total number of correct classification per class}}{\text{Row total}} \quad (1)$$

The PA is computed by dividing the number of correctly classified pixels in each land-use class by the number of reference pixels in that land use category (the column total)<sup>60</sup>. The PA represents how accurately reference pixels of land-use class are classified and the equation used to derive PA is given as Eq. 2.

$$PA = \frac{\text{Number of correctly classified pixels}}{\text{Column total}} \quad (2)$$

The overall accuracy (OA) assesses the match or mismatch between classified and ground truth of land use, and is derived by dividing total number of accurately classified pixels by total number of reference pixels<sup>50</sup>. The equation used to derived OA is given as Eq. 3.

$$\text{Overall accuracy} = \frac{\text{Number of pixels classified correctly}}{\text{Number of reference sampling pixels}} \quad (3)$$

The Kappa coefficient evaluates the reliability of classified raster which is calculated using error matrix. The coefficient values vary between 0 and 1, where a higher value means a better reliability and vice versa. The Kappa coefficient is calculated using Eq. 4<sup>50</sup>.

$$K = \frac{(\text{percent overall correct value} - \text{percent correct agreement to observed values})}{(\text{total number of class} - \text{percent correct agreement to observed values})} \quad (4)$$

F1-score or F-measure is a useful statistical approach to find the accuracy of classification using precision and recall<sup>61</sup>. F1-score is a harmonic mean of precision and recall, given in Eq. 5.

$$F1 - \text{score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

where precision, also known as the positive predictive value, is calculated by dividing the number of true positive values by the number of all (true and false) positive values (Eq. 6). Recall, also known as sensitivity, is calculated (Eq. 7) by dividing the number of true positive values by total the number of predicted values (true positives plus false negatives)<sup>62</sup>.

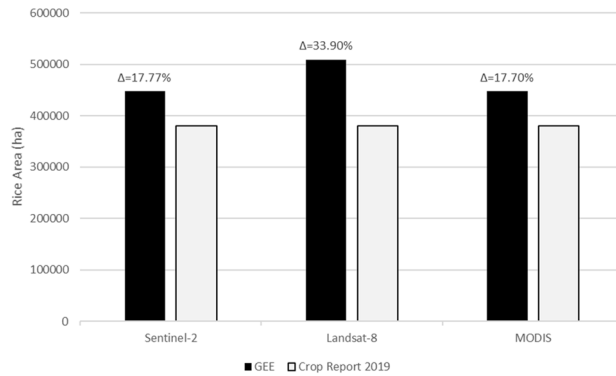
$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \quad (6)$$

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (7)$$

The classification process is only considered reliable if it meets accuracy checks, because land use land cover (LULC) maps derived from satellite images may contain some errors due to several factors ranging from techniques in classification to satellite based data retrieval methods<sup>63,64</sup>. For paddy rice maps (Fig. 8), a confusion matrix was developed containing kappa coefficient (K), F1-Score, UA, PA, and OA. It shows that OA was more than 90% for Sentinel-2 and Landsat-8 datasets which supports the reliability of derived maps (Table 2). For MODIS, OA was 82% showing least reliability for rice mapping compared to Sentinel-2 and Landsat-8.

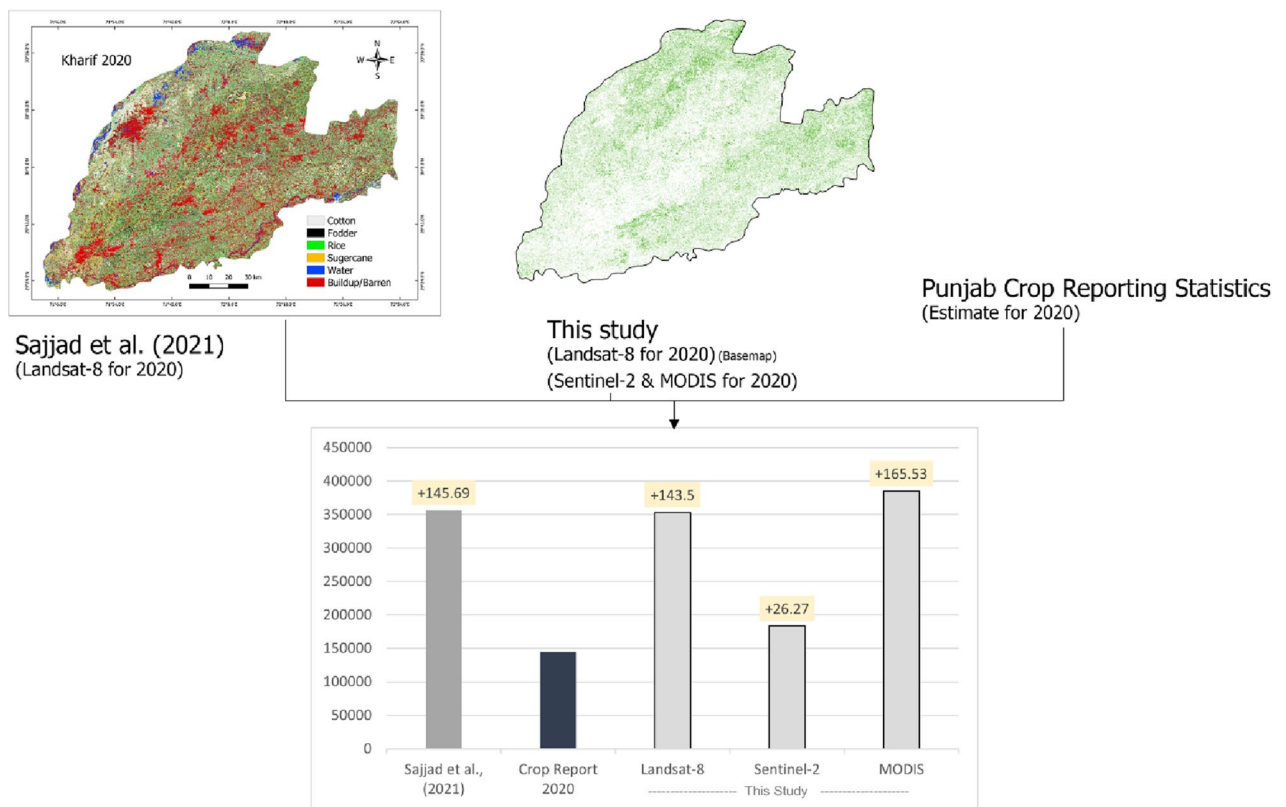
**Area estimation.** The area of rice crop is calculated from the classified maps using the “*ee.Image.pixelArea()*” function available inside GEE. This function computes and assign area for each pixel, and assign area values to each pixel. Then, we apply *ee.Reducer.sum* which computed total area per class of each classified rice map. For area comparison, crop reported area provided by Crop Reporting Service, Government of Punjab (<http://www.crs.agripunjab.gov.pk/reports>) was used as ground truth data to calculate mean difference between estimated and reported areas (Fig. 10). Multispectral instruments and ground truth data show a good agreement for estimated rice production. The mean percentage difference was approximately 17% for Sentinel-2 and MODIS while 33% for Landsat-8. Previous studies also reported overestimation<sup>37,65,66</sup>, which is mainly because of misclassified pixels in the rice class that may have contributed to some rise in the classified area.

For Landsat-8, the mean percentage difference was more significant (greater than 30%). This overestimation was due to higher misclassified rice pixels as observed previously through visual interpretation of Fig. 8.



**Figure 10.** Comparison of rice classified cultivated area from GEE and Crop statistics report (local agrarian data). The percentage of overestimation ( $\Delta$ ) by GEE classification relative to ground truth data is provided for each dataset.

Region: Multan Devison



**Figure 11.** Comparing our rice estimated area using three instruments with previously published study<sup>70</sup> and crop statistics report for 2020. Note: in the Figure area is compared for Multan division, and call-outs in bar plots shows the percentage increase or decrease in area compared to crop statistics report. The Figure is designed in Photoshop V23.3 software provided by Adobe (<https://www.adobe.com/products/photoshop>).

Surprisingly, the MODIS estimated area was nearly equal to the Sentinel-2 derived area; which reflects why MODIS is used by government officials and many researchers<sup>67–69</sup> for large-scale LULC area estimation. To further differentiate the performance of each instrument, and to address the issue why MODIS and Sentinel-2 derived paddy rice area was similar, we compared our study results (focusing on Multan division) with recently published study performed by Sajjad et al.<sup>70</sup> on crop classification in Multan. The Fig. 11 shows the comparison of paddy rice area from our study for Multan division with area published by Sajjad et al.<sup>70</sup> and Punjab crop statistics report for 2020. The Fig. 11 also shows the mean differences between different area sources compared with the base area estimate provided by Punjab crop statistics report for Multan region. From Fig. 11 it was observed that paddy rice area from Sajjad et al.<sup>70</sup> study using Landsat-8 also showed overestimation (+ 145.7%)

whereas our derived area using Landsat-8 also shows similar overestimation (+ 143.5%). Furthermore the Fig. 11 also highlighted a major difference in area estimation of Sentinel-2 (+ 26.3%) and MODIS (+ 165.5%) compared with Punjab crop statistics report.

## Discussion

Rice mapping is essential for bringing water-use efficiency, achieving regional food security, controlling disease transmission, and countering global warming. As our study area is a semi-arid region with lower precipitation (Fig. 5b), free available and GEE integrated MSI are suitable to map paddy rice crop due to lower cloud cover in the cropping months<sup>27,71,72</sup>. Despite their frequent use in mapping paddy fields, no study has compared the efficiency of Sentinel-2, Landsat-8 and MODIS datasets in classifying individual paddy rice fields. Our study compares these datasets by estimating their paddy crop classification potential, using accuracy assessment methods, and the mean difference between estimated area and area based on ground truth data.

The NDVI values derived from Sentinel-2 at 10 m spatial resolution (Fig. 5a) were used to track changes in crop growth stages. The NDVI of any crop differs spatially and temporally and thus can be used as a valuable tool to study crop growth changes over space and time<sup>48,49,71,73,74</sup>. The temporal variation shows that rice sowing and transplanting stages start from May until June in southern Punjab region. Lower NDVI values (Fig. 5a) in July and August reflect significant NIR adsorption on the water surfaces which in this case are paddy rice fields flooded with water and with minimum green area. Similarly, Usman et al.<sup>50</sup> stated that the individual crops starting time and crop-cycle length can be visualized easily from NDVI trends. According to them, low positive values of NDVI (0.1 or less) represent barren areas of rock, sand, or snow, while 0.1–0.2 represent soils. Most vegetation has moderate NDVI values (~0.2 to 0.5), while dense forests show high NDVI values (~0.6 to 0.9). The class 'rice' exhibits a unique trend compared to other crops (Cotton, and Sugarcane) in the Kharif season (Fig. 5a). For rice crop, the initial NDVI variation is a bit slow and lengthy because of rice nursery growth in June. However, the latter part of NDVI variation reaches its maximum due to rapid rice crop growth in September and October. Figure 7b shows variation in temperature and precipitation patterns during the Kharif season. It was observed that from July to September, mean precipitation values were high along with moderate LST (Day and Night), which supported rice cultivation<sup>14,52</sup>.

For rice crop, Sentinel-2 and Landsat-8 data showed similar NDVI trends (Fig. 4). Due to their high spatial and temporal resolution, these trends were successful in differentiating various rice growth stages. Conversely, the MODIS curve showed a peak response in August, mainly due to its coarse resolution (of 250 m). Although MODIS has a higher revisit time (1 day) than Sentinel-2 (5 days) and Landsat-8 (16 days), it cannot be used to study the growth cycle of individual rice fields due to its coarser resolution. Likewise Chen et al.<sup>75</sup> concluded that the MODIS imagery with 250 m resolution could not provide accurate vegetation information on corn growth. The variation in Sentinel-1 backscatter coefficient for paddy rice (Fig. 4) showed that the coefficient is very sensitive to water in paddy rice fields. The backscatter values were low in June and July due to the presence of water in the fields, and were high in September and October due to high vegetation. Singha et al.<sup>6</sup> noted a similar trend in the the backscatter values. The Sentinel-1 VH backscatter coefficient can be used as an indicator for tracking paddy rice growth as the backscatter values change with the varying conditions of paddy rice stages. Unlike other crops, rice spends a substantial time submerged in water. Thus, it can be easily differentiated from other crops. Also, it has distinct growth stages, which can be observed using the Sentinel-1 VH backscatter coefficient. Based on the temporal variation of NDVI and Sentinel-1 backscatter coefficient, an accurate rice growth profile was generated for the Southern Punjab region for the year 2020, which was used as a baseline in rice fields sampling. The coordinates used for rice training and validation by global positioning system (GPS) were crosschecked using NDVI and backscatter temporal variation. This step was performed inside GEE to ensure that only rice fields coordinates having the same trend as observed in Fig. 5 were taken for further classification and accuracy assessment. Previous studies<sup>6,20,76–78</sup> support the RF classifier's capability in identifying paddy fields. Therefore, this study utilized RF classifier for the classification of paddy rice.

Results showed that Sentinel-2 performed best in classifying individual rice fields (Figs. 8, 9). MODIS failed to accurately classify rice fields, and it formed clusters with neighboring fields. A visual interpretation was performed for comparing the efficiency of each dataset in classifying individual rice fields. It was observed that Landsat-8 efficiently differentiated paddy rice fields; however, it misclassified rice pixels at some points with other crops having similar vegetation characteristics. Overall, Sentinel-2 performed best in accuracy assessment than both Landsat-8 and MODIS (Table 2).

Further, we compare estimated and reported areas for rice production, which shows an overestimation of ~18% for both Sentinel-2 and MODIS, and ~34% overestimation for Landsat-8 (Fig. 10). Similar overestimation was reported by other researchers<sup>37,65,66</sup>. For example, Mananze et al.<sup>37</sup> used GEE to map shifting agriculture dynamics in Mozambique and found 15.5% of mean difference between the estimated and ground truth areas for agricultural land use. In another study to map cropping intensity of small-scale farms in India, Jain et al.<sup>79</sup> reported a large discrepancy (with percentage difference of 150.8%) between cropland area estimated through remote sensing (Landsat data) and the area obtained from agricultural census. Likewise, Dheeravath et al.<sup>80</sup> found that irrigation area derived using MODIS (500 m) data was greater than the ground truth data (with the mean percentage difference of 17.2%) in most Indian States.

To compare the overestimation in paddy rice area, we compared our results with a previous study performed by Sajjad et al.<sup>70</sup> focusing on Multan division (Fig. 11). The Fig. 11 shows area estimated as well as mean percentage difference of area from each source compared with crop statistics report. It was observed that there exist a similarity in overestimation of Landsat-8 derived area reported by Sajjad et al. (2021) (145.7%) and our Landsat-8 derived paddy rice area for Multan division (143.5%). The Punjab crop statistics report provided area was taken as base area for comparison. When compared, Sentinel-2 with mean percentage overestimation of 26.3%

stands above Landsat-8 and MODIS with overestimation 143.5% and 165.5% respectively. To conclude, MODIS performance was least, as this instrument overestimated the area for Multan division with an overestimation of 165.5%, far above the percentage difference of Landsat-8. This reflects why for crop area estimation MODIS should be avoided and preference should be towards Sentinel-2.

The multispectral imagery-based paddy rice mapping implemented in GEE includes several benefits. First, there is reduction in data acquiring and pre-processing time<sup>40</sup>. Otherwise, it might have taken longer to download and pre-process the datasets for a large study area like ours (i.e., southern Punjab). Second, there is a significant decrease in computational time<sup>81</sup>. The MSI-based RF algorithm inside the GEE performed paddy rice fields classification in a very short time. Thus, high-performance computing resources like GEE facilitate quick and rapid mapping of paddy rice planting areas at a large scale<sup>4,6</sup>. Future studies of paddy rice mapping under similar climatic conditions can utilize the most efficient dataset found in this study: Sentinel-2. In addition, further research may focus on rice mapping for the entire country using high-resolution Sentinel-2 imagery with a similar GEE-based methodology.

## Conclusions

In this study, we compared three, free and open, satellite datasets (Sentinel-2 (10 m), Landsat-8 (30 m), and MODIS (250 m)) to evaluate their efficiency/effectivity for mapping of rice crop in southern Punjab, Pakistan. The results show that rice classification improves with increase in spatial resolution. Among the three, the coarsest resolution dataset, MODIS, do not precisely classify individual rice fields. However, the other two high-resolution datasets (i.e., Sentinel-2 and Landsat-8) can accurately track crop growth stages for making its growth profiles and studying temporal trends. From the visual interpretation, accuracy assessment, and mean difference in area estimation, we observe that MODIS was not able to classify individual rice fields correctly in this case. However, Landsat-8 classified the paddy rice fields accurately with moderate misclassification. In comparison to MODIS and Landsat-8, Sentinel-2 performed best in differentiating individual rice fields with lower misclassification and showed a lesser mean difference in area estimation. Due to limited resources and training data, this study focused only on the RF classifier to classify southern Punjab divisions. JavaScript API is used inside GEE, which facilitates future studies to implement the same methodology in GEE with ease. Future studies may focus on (1) rice mapping of the whole of Pakistan, which could be easily implemented inside GEE using the same methodology and code, (2) comparison of different machine learning classifiers (like Random Forest RF, Support Vector Machine, and Artificial Neural Network to classify paddy rice efficiently, and (3) comparison of Synthetic Aperture Radar and Multispectral Instruments to map paddy rice efficiently.

## Data availability

All the datasets used in this study are open access, and references are provided at first mention. Paddy rice classification images for Southern Punjab are also available, and can be provided upon reasonable request to corresponding authors. No direct plant material is used in this study. Only open access multispectral satellite images are used, that are freely available inside GEE data catalogue (<https://www.developers.google.com/earth-engine/datasets>).

Received: 13 February 2022; Accepted: 26 July 2022

Published online: 01 August 2022

## References

- Dong, J. *et al.* Mapping paddy rice planting area in northeastern Asia with Landsat 8 images, phenology-based algorithm and Google Earth Engine. *Remote Sens. Environ.* **185**, 142–154. <https://doi.org/10.1016/j.rse.2016.02.016> (2016).
- GOP, G. J. *Economic Advisor's Wing: Islamabad, Pakistan*. Pakistan Economic Survey 2017–18. Finance Division (2018).
- Luo, T., Young, R. & Reig, P. J. *Aqueduct Projected Water Stress Country Rankings*. Technical Note (2015).
- Nizami, A., Zulfiqar, M., Ali, J., Khan, N. & Sheikh, I. Improving water productivity in rice—A response to climate change and water stress in Pakistan. *Sarhad J. Agric.* <https://doi.org/10.17582/journal.sja/2020/36.2.383.388> (2020).
- Sporchia, F., Thomsen, M. & Caro, D. J. Drivers and trade-offs of multiple environmental stressors from global rice. *Sustain. Prod. Consum.* **26**, 16–32 (2020).
- Singha, M., Dong, J., Zhang, G. & Xiao, X. High resolution paddy rice maps in cloud-prone Bangladesh and Northeast India using Sentinel-1 data. *Sci. Data* **6**, 26. <https://doi.org/10.1038/s41597-019-0036-3> (2019).
- Segarra, J., Buchaillet, M. L., Araus, J. L. & Kefauver, S. C. Remote sensing for precision agriculture: Sentinel-2 improved features and applications. *Agronomy* **10**, 641 (2020).
- Reyes-González, A., Kjaersgaard, J., Trooien, T., Hay, C. & Ahiablame, L. Estimation of crop evapotranspiration using satellite remote sensing-based vegetation index. *Adv. Meteorol.* **1–12**, 2018. <https://doi.org/10.1155/2018/4525021> (2018).
- Matese, A. *et al.* Estimation of water stress in grapevines using proximal and remote sensing methods. *Remote Sens.* <https://doi.org/10.3390/rs10010114> (2018).
- Fritz, S. *et al.* A comparison of global agricultural monitoring systems and current gaps. *Agric. Syst.* **168**, 258–272 (2019).
- Shanmugapriya, P., Rathika, S., Ramesh, T. & Janaki, P. Applications of remote sensing in agriculture—A review. *Int. J. Curr. Microbiol. Appl. Sci.* **8**, 2270–2283. <https://doi.org/10.20546/ijcmas.2019.801.238> (2019).
- Mehboob, S. R. *Governance and Militancy in Pakistan's Southern Punjab Region*. <https://www.csis.org/analysis/governance-and-militancy-pakistan%E2%80%99s-southern-punjab-region> (2011).
- Aslam, A. Q. *et al.* Integrated climate change risk assessment and evaluation of adaptation perspective in southern Punjab, Pakistan. *Sci. Total Environ.* **628–629**, 1422–1436. <https://doi.org/10.1016/j.scitotenv.2018.02.129> (2018).
- Chandio, A. A., Magsi, H. & Ozturk, I. Examining the effects of climate change on rice production: Case study of Pakistan. *Environ. Sci. Pollut. Res. Int.* **27**, 7812–7822. <https://doi.org/10.1007/s11356-019-07486-9> (2020).
- Saadat, M., Hasanlou, M. & Homayouni, S. Rice crop mapping using Sentinel-1 time series images (Case study: Mazandaran, Iran). *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **18**, 897–904. <https://doi.org/10.5194/isprs-archives-XLII-4-W18-897-2019> (2019).
- Gorelick, N. *et al.* Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **202**(202), 18–27 (2017).

17. Kumar, L. & Mutanga, O. Google Earth engine applications since inception: usage, trends, and potential. *Remote Sens.* <https://doi.org/10.3390/rs10101509> (2018).
18. Chong, L. *et al.* Mapping the fallowed area of paddy fields on Sanjiang Plain of Northeast China to assist water security assessments. *J. Integr. Agric.* **19**, 1885–1896 (2020).
19. Wu, W., Wang, W., Meadows, M. E., Yao, X. & Peng, W. Cloud-based typhoon-derived paddy rice flooding and lodging detection using multi-temporal Sentinel-1&2. *Front. Earth Sci.* **13**, 682–694 (2019).
20. Wang, Y., Zang, S. & Tian, Y. Mapping paddy rice with the random forest algorithm using MODIS and SMAP time series. *Chaos Solitons Fractals* <https://doi.org/10.1016/j.chaos.2020.110116> (2020).
21. Cai, Y., Lin, H. & Zhang, M. Mapping paddy rice by the object-based random forest method using time series Sentinel-1/Sentinel-2 data. *Adv. Space Res.* **64**, 2233–2244 (2019).
22. Zhang, M., Lin, H., Wang, G., Sun, H. & Fu, J. Mapping paddy rice using a convolutional neural network (CNN) with Landsat 8 datasets in the Dongting Lake Area, China. *Remote Sens.* **10**, 1840 (2018).
23. Zhang, G. *et al.* Spatiotemporal patterns of paddy rice croplands in China and India from 2000 to 2015. *Sci. Total Environ.* **579**, 82–92. <https://doi.org/10.1016/j.scitotenv.2016.10.223> (2017).
24. Hassan, G. & Hassan, F. J. E. W. Sustainable use of groundwater for irrigated agriculture: A case study of Punjab. *Pakistan.* **57**, 475–480 (2017).
25. Dong, J. & Xiao, X. Evolution of regional to global paddy rice mapping methods: A review. *J. Photogramm. Remote Sens.* **119**, 214–227 (2016).
26. Qiu, B., Lu, D., Tang, Z., Chen, C. & Zou, F. Automatic and adaptive paddy rice mapping using Landsat images: Case study in Songnen Plain in Northeast China. *Sci. Total Environ.* **598**, 581–592 (2017).
27. Jiang, Y. *et al.* Large-scale and high-resolution crop mapping in China using Sentinel-2 satellite imagery. *Agriculture* **10**, 433 (2020).
28. Immitzer, M., Vuolo, F. & Atzberger, C. J. First experience with Sentinel-2 data for crop and tree species classifications in central Europe. *Remote Sens.* **8**, 166 (2016).
29. Sentinel, E. *3-Missions-Sentinel Online* (ESA, Paris, 2014).
30. Adiri, Z., Lhissou, R., El Harti, A., Jellouli, A. & Chakouri, M. Recent advances in the use of public domain satellite imagery for mineral exploration: A review of Landsat-8 and Sentinel-2 applications. *Ore Geol. Rev.* **117**, 103332 (2020).
31. Mohajane, M. *et al.* Land use/land cover (LULC) using landsat data series (MSS, TM, ETM+ and OLI) in Azrou Forest, in the Central Middle Atlas of Morocco. *Environments* **5**, 131 (2018).
32. García-Mora, T. J., Mas, J.-F. & Hinkley, E. A. Land cover mapping applications with MODIS: a literature review. *Int. J. Digit. Earth* **5**, 63–87 (2012).
33. French, A. N. *et al.* Satellite-based NDVI crop coefficients and evapotranspiration with eddy covariance validation for multiple durum wheat fields in the US Southwest. *Agric. Water Manag.* **239**, 106266 (2020).
34. Morin, G., Roux, R. L., Lemasle, P.-G. & Quénoel, H. Mapping bioclimatic indices by downscaling MODIS land surface temperature: Case study of the Saint-Emilion area. *Remote Sens.* **13**, 4 (2021).
35. Ullah, W. *et al.* Large-scale atmospheric circulation patterns associated with extreme monsoon precipitation in Pakistan during. *Atmos. Res.* **1981–2018**, 105489 (2021).
36. Tian, F., Wu, B., Zeng, H., Zhang, X. & Xu, J. Efficient identification of corn cultivation area with multitemporal synthetic aperture radar and optical images in the Google Earth Engine Cloud Platform. *Remote Sens.* <https://doi.org/10.3390/rs11060629> (2019).
37. Mananze, S., Pôças, I. & Cunha, M. Mapping and assessing the dynamics of shifting agricultural landscapes using Google Earth Engine cloud computing, a case study in Mozambique. *Remote Sens.* <https://doi.org/10.3390/rs12081279> (2020).
38. Kumar, L. & Mutanga, O. Google Earth Engine applications since inception: Usage, trends, and potential. *Remote Sens.* **10**, 1509 (2018).
39. Huang, H. *et al.* Mapping major land cover dynamics in Beijing using all Landsat images in Google Earth Engine. *Remote Sens.* **202**, 166–176 (2017).
40. Amani, M. *et al.* Google Earth Engine cloud computing platform for remote sensing big data applications: A comprehensive review. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **13**, 5326–5350. <https://doi.org/10.1109/jstars.2020.3021052> (2020).
41. Tamiminia, H. *et al.* Google Earth Engine for geo-big data applications: A meta-analysis and systematic review. *ISPRS J. Photogramm. Remote Sens.* **164**, 152–170. <https://doi.org/10.1016/j.isprsjprs.2020.04.001> (2020).
42. Maxwell, A. E., Warner, T. A. & Fang, F. Implementation of machine-learning classification in remote sensing: An applied review. *Int. J. Remote Sens.* **39**, 2784–2817 (2018).
43. Shakoor, U., Saboor, A., Baig, I., Afzal, A. & Rahman, A. Climate variability impacts on rice crop production in Pakistan. *Pak. J. Agric. Res.* **28**(1), 19–27 (2015).
44. Memon, N. A. Rice: Important cash crop of Pakistan. *Pak. Food J* **26**(7), 21–23 (2013).
45. FAO. *FAO Statistical Databases (Food and Agriculture Organization of the United Nations) Databases.* <http://digital.library.wisc.edu/1711.web/faostat> (2020).
46. Bellón, B., Bégué, A., Lo Seen, D., De Almeida, C. A. & Simões, M. A remote sensing approach for regional-scale mapping of agricultural land-use systems based on NDVI time series. *Remote Sens.* **9**, 600 (2017).
47. Cao, R. *et al.* A simple method to improve the quality of NDVI time-series data by integrating spatiotemporal information with the Savitzky-Golay filter. *Remote Sens.* **217**, 244–257 (2018).
48. Shammi, S. A. & Meng, Q. Use time series NDVI and EVI to develop dynamic crop growth metrics for yield modeling. *Ecol. Indic.* **121**, 107124 (2020).
49. Tong, X. *et al.* Revisiting the coupling between NDVI trends and cropland changes in the Sahel drylands: A case study in western Niger. *Remote Sens. Environ.* **191**, 286–296 (2017).
50. Usman, M., Liedl, R., Shahid, M. & Abbas, A. Land use/land cover classification and its change detection using multi-temporal MODIS NDVI data. *J. Geogr. Sci.* **25**, 1479–1506 (2015).
51. Funk, C. *et al.* The climate hazards infrared precipitation with stations—A new environmental record for monitoring extremes. *Sci. Data* **2**, 1–21 (2015).
52. Boonwichai, S., Shrestha, S., Babel, M. S., Weesakul, S. & Datta, A. Climate change impacts on irrigation water requirement, crop water productivity and rice yield in the Songkhram River Basin, Thailand. *J. Clean. Prod.* **198**, 1157–1164. <https://doi.org/10.1016/j.jclepro.2018.07.146> (2018).
53. Waleed, M. & Sajjad, M. Leveraging cloud-based computing and spatial modeling approaches for land surface temperature disparities in response to land cover change: Evidence from Pakistan. *Remote Sens. Appl. Soc. Environ.* **25**, 100665 (2022).
54. Mateo-García, G., Gómez-Chova, L., Amorós-López, J., Muñoz-Mari, J. & Camps-Valls, G. J. Multitemporal cloud masking in the Google Earth Engine. *Remote Sens.* **10**, 1079 (2018).
55. Song, X.-P., Huang, W., Hansen, M. C. & Potapov, P. An evaluation of Landsat, Sentinel-2, Sentinel-1 and MODIS data for crop type mapping. *Sci. Remote Sens.* **3**, 100018. <https://doi.org/10.1016/j.srs.2021.100018> (2021).
56. Xiong, J. *et al.* Automated cropland mapping of continental Africa using Google Earth Engine cloud computing. *ISPRS J. Photogramm. Remote Sens.* **126**, 225–244 (2017).
57. Kuenzer, C. & Knauer, K. Remote sensing of rice crop areas. *Int. J. Remote Sens.* **34**, 2101–2139. <https://doi.org/10.1080/01431161.2012.738946> (2012).

58. Lu, D., Li, G., Moran, E. & Hetrick, S. Spatiotemporal analysis of land use and land cover change in the Brazilian Amazon. *Int. J. Remote Sens.* **34**, 5953–5978. <https://doi.org/10.1080/01431161.2013.802825> (2013).
59. Rwanga, S. S. & Ndambuki, J. M. Accuracy assessment of land use/land cover classification using remote sensing and GIS. *Int. J. Geosci.* **8**, 611 (2017).
60. Lillesand, T., Kiefer, R. W. & Chipman, J. *Remote Sensing and Image Interpretation* (Wiley, Hoboken, 2015).
61. Tharwat, A. Classification assessment methods. *Appl. Comput. Inform.* **17**, 168–192 (2020).
62. Mahdianpari, M., Salehi, B., Rezaee, M., Mohammadimanesh, F. & Zhang, Y. Very deep convolutional neural networks for complex land cover mapping using multispectral remote sensing imagery. *Remote Sens.* **10**, 1119 (2018).
63. Hussain, S. *et al.* Using GIS tools to detect the land use/land cover changes during forty years in Lodhran District of Pakistan. *Environ. Sci. Pollut. Res. Int.* **27**, 39676–39692. <https://doi.org/10.1007/s11356-019-06072-3> (2020).
64. Hussain, S. *et al.* Study of land cover/land use changes using RS and GIS: a case study of Multan district, Pakistan. *Environ. Monit. Assess.* **192**, 2. <https://doi.org/10.1007/s10661-019-7959-1> (2019).
65. Gusso, A., Formaggio, A. R., Rizzi, R., Adami, M. & Rudorff, B. F. T. J. P. A. B. Soybean crop area estimation by Modis/Evi data. *Pesqui. Agropecu. Bras.* **47**, 425–435 (2012).
66. Zhang, H. *et al.* Crop classification and acreage estimation in North Korea using phenology features. *GISci. Remote Sens.* **54**, 381–406. <https://doi.org/10.1080/15481603.2016.1276255> (2017).
67. Seydi, S. T., Akhoondzadeh, M., Amani, M. & Mahdavi, S. Wildfire damage assessment over Australia using Sentinel-2 imagery and MODIS land cover product within the Google Earth Engine cloud platform. *Remote Sens.* **13**, 220 (2021).
68. Bansal, S., Katyal, D. & Garg, J. A novel strategy for wetland area extraction using multispectral MODIS data. *Remote Sens. Environ.* **200**, 183–205 (2017).
69. Chen, Y. *et al.* Mapping croplands, cropping patterns, and crop types using MODIS time-series data. *Int. J. Appl. Earth Obs. Geoinform.* **69**, 133–147 (2018).
70. Sajjad, H. *et al.* Satellite-based evaluation of temporal change in cultivated land in Southern Punjab (Multan region) through dynamics of vegetation and land surface temperature. *Open Geosci.* **13**, 1561–1577. <https://doi.org/10.1515/geo-2020-0298> (2021).
71. Ghosh, P., Mandal, D., Bhattacharya, A., Nanda, M. K. & Bera, S. Assessing crop monitoring potential of Sentinel-2 in a spatio-temporal scale. *Int. Arch. Photogr. Remote Sens. Spat. Inf. Sci.* **425**, 227–231 (2018).
72. Phiri, D. *et al.* Sentinel-2 data for land cover/use mapping: A review. *Remote Sens.* **12**, 2291 (2020).
73. Veloso, A. *et al.* Understanding the temporal behavior of crops using Sentinel-1 and Sentinel-2-like data for agricultural applications. *Remote Sens. Environ.* **199**, 415–426 (2017).
74. Ke, Y., Im, J., Lee, J., Gong, H. & Ryu, Y. Characteristics of Landsat 8 OLI-derived NDVI by comparison with multiple satellite sensors and in-situ observations. *Remote Sens. Environ.* **164**, 298–313 (2015).
75. Chen, P.-Y., Fedosejevs, G., Tiscareno-Lopez, M. & Arnold, J. G. Assessment of MODIS-EVI, MODIS-NDVI and VEGETATION-NDVI composite data using agricultural measurements: An example at corn fields in western Mexico. *Environ. Monit. Assess.* **119**, 69–82 (2006).
76. Tatsumi, K., Yamashiki, Y., Torres, M. A. C. & Taipe, C. L. R. Crop classification of upland fields using Random forest of time-series Landsat 7 ETM+ data. *Comput. Electron. Agric.* **115**, 171–179 (2015).
77. Pal, M. Random forest classifier for remote sensing classification. *Int. J. Remote Sens.* **26**, 217–222. <https://doi.org/10.1080/0143160412331269698> (2007).
78. Hao, P., Zhan, Y., Wang, L., Niu, Z. & Shakir, M. J. R. S. Feature selection of time series MODIS data for early crop classification using random forest: A case study in Kansas, USA. *Remote Sens.* **7**, 5347–5369 (2015).
79. Jain, M., Mondal, P., DeFries, R. S., Small, C. & Galford, G. L. Mapping cropping intensity of smallholder farms: A comparison of methods using multiple sensors. *Remote Sens. Environ.* **134**, 210–223 (2013).
80. Dheeravath, V. *et al.* Irrigated areas of India derived using MODIS 500 m time series for the years 2001–2003. *ISPRS J. Photogramm. Remote Sens.* **65**, 42–59 (2010).
81. Gorelick, N. *et al.* Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **202**, 18–27. <https://doi.org/10.1016/j.rse.2017.06.031> (2017).

## Author contributions

Conceptualization: M.W., M.M.; M.H.R. Methodology: A.A., A.A.; Software: H.U.F., S.H., M.H.R.; Formal analysis and results: M.A., S.A.Q.; Writing—original draft preparation: M.W., M.M.; Writing—review and editing: W.N., M.H.R., H.M.R.J., N.M., T.A., F.M.; Funding acquisition: A.ELS., M.H.R.

## Funding

Open Access funding enabled and organized by Projekt DEAL.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to M.M., M.H.-u. or A.E.S.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022