Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Wheat Mapping Using Multitemporal Sentinel 2 Imagery and Machine Learning Classification in Multan District, Punjab Pakistan

Hassan Raza

hraza.ms26igis@student.nust.edu.pk

National University of Sciences and Technology

Hafsa Naveed

University of the Punjab

Muhammad Sarfraz

University of the Punjab

Shahid Parvez

University of the Punjab

Saira Bano

University of the Punjab

Research Article

Keywords: Wheat mapping, Sentinel-2, support vector machine, vegetation indices, crop phenology, Google Earth Engine

Posted Date: November 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11578443/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Wheat Mapping Using Multitemporal Sentinel 2 Imagery and Machine Learning Classification in Multan District, Punjab Pakistan

***Hassan Raza[1,3], Hafsa Naveed[1,2], Muhammad Sarfraz[1], Shahid Parvez[1], Saira Bano[2]**

[1]*Institute of Space Science, University of the Punjab, Lahore, Pakistan*

[2]*College of Earth & Environmental Science, University of the Punjab, Lahore, Pakistan*

[3]*Institute of Geographical Information System, National University of Science & Technology, Islamabad, Pakistan*

[*]*Corresponding author: hraza.ms26igis@student.nust.edu.pk*

Highlights

What are the key results?

- Wheat mapping was performed using SVM-RBF with 54 bands calibrated with gamma = 1/54 ≈ 0.019, which gives an Overall Accuracy of 93.3% compared to 87.5% and 86.6% obtained by Random Forest and CART, respectively, on the same sample partition in Multan District.
- LSWI collapses 10–14 days earlier than NDVI and is thus not only an earlier indicator of collapse, but also has operational value in terms of advance warning of a harvest.

What are the implications of the key findings?

- It is important to calibrate the gamma hyperparameter for high-dimensional SVM-RBF crop classification to 1/n features and such calibration should be considered as standard practice for multi-index Sentinel-2 feature stacks.
- The fully cloud-based and open data Google Earth Engine (GEE) workflow is readily scalable to the provincial level, cost-effective, and complementary to Pakistan's traditional girdawari survey system.

Abstract: In Pakistan, an accurate and timely estimation of wheat cultivation is essential for food security planning and agricultural monitoring. In this study, we developed a multi-temporal remote sensing approach for wheat mapping in Multan District during the Rabi growing season 2024-2025. Three complementary vegetation indices (NDVI, EVI, and LSWI) were derived from 18 consecutive 10-day composite periods of 127 cloud-screened Sentinel-2 Multi Spectral Imager (MSI) Level-2A scenes, resulting in a 54-band feature stack for each composite period. A Support Vector Machine (SVM) with a calibrated Radial Basis Function kernel (gamma = 1/54 ≈ 0.019) achieved an Overall Accuracy (OA) of 93.3% and kappa coefficient of 0.865 for 270 validation samples, outperforming the Random Forest (OA: 91.8%, Kappa: 0.835) and CART (OA: 81.5%, Kappa: 0.627) classifiers. The wheat area was estimated at 457,837 acres (185,274 ha), within +4.5% of the official Government of Punjab figure. Phenological analysis revealed that the LSWI undergoes an early post-anthesis collapse 10–14 days ahead of the NDVI, providing an operationally useful advance harvest indicator. This study delivers a 10-metre-resolution wheat map for the Multan District for the complete

2024–25 Rabi season and demonstrates a cloud-based, open-data GEE workflow scalable to provincial-level crop monitoring in Pakistan.

Keywords: Wheat mapping; Sentinel-2; support vector machine; vegetation indices; crop phenology; Google Earth Engine

1. Introduction

Food security is among the foremost challenges of the twenty-first century. The United Nations projects a global population of 9.7 billion by 2050, necessitating more efficient and data-driven agricultural production systems [1,2]. Wheat (Triticum aestivum L.) is the most strategically important food crop globally, providing more than 20% of the total caloric intake and serving as the primary protein source for an estimated 2.5 billion people [3,4]. Wheat has a unique importance in Pakistan; it is used in about 72% of the per capita consumption of cereal, utilized about 8.9 million hectares of cultivated land every year, and directly or indirectly involves the labor of 60–65% of the rural population [5].

According to official estimates of the Government of Punjab, the wheat sown area in Multan for Rabi 2024–25 was 438,000 acres (177,400 ha) [6]. As important as this is, Pakistan's chief crop monitoring system, the girdawari administrative survey by the village level revenue officers (patwaris) suffers from inter-assessor variability, administrative pressures and lack of spatial explicitness [7,8]. These restrictions hinder government decision-making on procurement, import/export policies, and interventions in the field of food security in based on evidence.

Multi-temporal satellite remote sensing is a novel approach to traditional survey techniques that involves machine learning classification and cloud computing. The European Space Agency (ESA) under the Copernicus programe has launched a 10-metre-resolution multispectral image acquired every 5 days with no cost, and the Sentinel-2 Multispectral Instrument (MSI) is ideally suited to operate as an operational crop monitoring instrument in data-scarce situations. In addition, the Google Earth Engine (GEE) cloud platform further democratizes large-scale geospatial analysis without requiring institutional HPCs [9,10].

There is a significant amount of literature pertaining to remote sensing, which supports the use of multitemporal optical imagery as a good foundation for crop type mapping. Weiss et al. [11] reviewed how agricultural remote sensing evolved from single-date classification to dense multi-temporal time series, which can capture the phenology of the entire crop. Wardlow et al. [12] found that using time-series vegetation index profiles significantly improves the

discrimination of crops from large agricultural land areas, consistent with the results of Immitzer et al. [13] who illustrated that Sentinel-2's spatio-spectral resolution offers good crop and tree species separability, even from a single scene, and that temporal stacking could further enhance discrimination. This benefit can be quantified directly, as reported by Vuolo et al. [14] who found that multi-temporal Sentinel-2 feature stacks increased crop classification performance by 5-10 percentage points compared with single-date classification for various test locations in Europe [15]. These findings collectively motivate the multi-temporal, multi-index feature engineering strategy adopted in the present study and underline the continued need for locally calibrated, machine-learning-based wheat mapping studies in South Asian irrigated systems such as Multan District, where crop calendars, field sizes, and atmospheric conditions differ substantially from the temperate and semi-arid systems in which much of the prior literature was developed [16].

Support Vector Machine (SVM) classifiers, introduced by Cortes and Vapnik [17] and extended with kernel methods by Burges [18,19], have consistently achieved state-of-the-art accuracy for crop mapping using multitemporal satellite imagery. Comparative studies in South Asia have reported overall accuracies exceeding 90% for wheat classification using Sentinel-2 and SVM [20,21,22]. However, prior wheat mapping studies in Pakistan, such as those by Qureshi et al. [20] and Usman et al. [21], have relied principally on single-index (NDVI-only) feature stacks and have not systematically addressed the gamma hyperparameter calibration required for high-dimensional SVM with a Radial Basis Function (SVM-RBF) kernel classifiers. SVM was selected as the primary classifier for this study because its maximum-margin formulation is theoretically well-suited to high-dimensional, moderate-sample-size feature spaces such as the 54-band multi-temporal stack used here, and because it has consistently outperformed alternative classifiers such as Random Forest and CART in comparable crop-mapping applications. This choice is empirically validated in Section 3.1, through a direct comparison against Random Forest and CART on identical sample partitions. This study explicitly fills these methodological gaps.

The specific contributions of this study are as follows: (i) a 10-metre-resolution wheat map of Multan District for the complete 2024–25 Rabi season; (ii) a quantified comparison of machine learning classifier accuracy across SVM, Random Forest, and CART classifiers on an identical training/validation partition; (iii) a multi-index (NDVI+EVI+LSWI) feature engineering approach with documented phenological interpretation; and (iv) validation of the estimated

wheat area against official Punjab Agriculture Department statistics at both district and sub-district (tehsil) levels.

2. Materials and Methods

2.1. Study Area

Multan District is located in South-Central Punjab, Pakistan (71°04'–71°52'E longitude, 29°22'–30°28 'N latitude), encompassing a total area of approximately 3,721 km². The district is comprised of four tehsils: Multan City, Multan Saddar, Shujabad, and Jalalpur Pirwala. The climate is semi-arid and hot summer (December – January minimum temperature 4 – 8°C), and the Rabi wheat system in the canal irrigated was the main agricultural activity due to below-normal winter rainfall during the growing season 2024 – 25. The mean annual rainfall is approximately 200 mm, and the Lower Bari Doab Canal system provides irrigation. Multan has been identified as one of the most productive wheat-producing districts of Punjab [6] and in wheat production statistics, it is always included in the top five wheat-producing districts of Punjab.

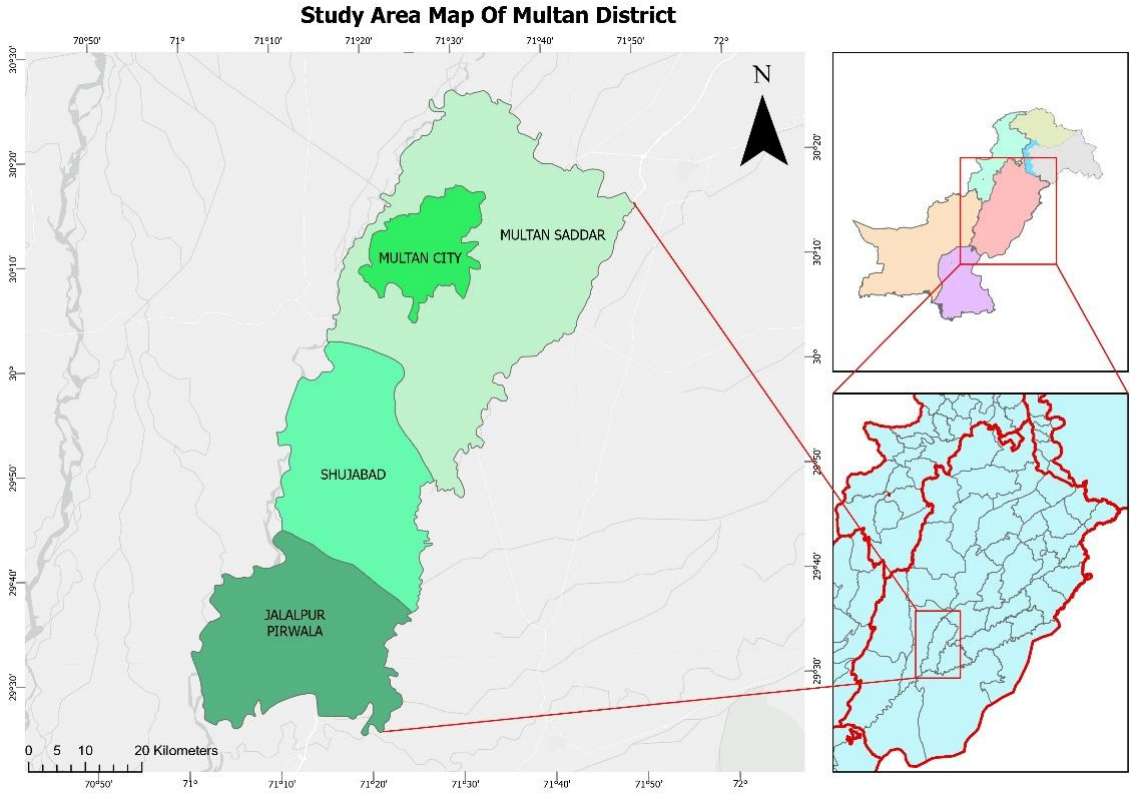


Figure 1. *Study area map of Multan District, Southern Punjab, Pakistan (29.22–30.28°N, 71.07–71.87°E), showing the four administrative tehsils (Multan City, Multan Saddar, Shujabad, and Jalalpur Pirwala).*

2.2. Satellite Data and Preprocessing

Sentinel-2 MSI Level-2A (Surface Reflectance) data were acquired for the 2024-25 Rabi growing season (Nov 24-April 25) from the GEE data catalogue. The 127 scenes passed the initial cloud screening filter (QA60 cloud mask, cloud fraction (CF) < 30% at the scene level). This count incorporates the combined twin-satellite Sentinel-2A/2 B constellation, which decreases the effective revisit interval down to approximately 2-3 days, as well as the overlap between two neighboring tile footprints that cover Multan District, resulting in significantly more scenes available than the nominal 5-day revisit time for Sentinel-2A/2 B would indicate over the 182-day study window (127 were cloud-screened). Three vegetation indices were computed for each scene from the following Sentinel-2 bands: band 4 (red, 665 nm, 10 m), band 8 (NIR, 842 nm, 10 m), band 2 (blue, 490 nm, 10 m), and band 11 (SWIR-1, 1610 nm, 20 m; resampled to 10 m using bilinear interpolation):

NDVI is defined below, together with EVI and LSWI (Equations 1–3):

$$NDVI = \frac{Band\,8 - Band\,4}{Band\,8 + Band\,4}$$

$$EVI = 2.5 \times \frac{Band\,8 - Band\,4}{Band\,8 + 6 \times Band\,4 - 7.5 \times Band\,2 + 1}$$

$$LSWI = \frac{Band\,8 - Band\,11}{Band\,8 + Band\,11}$$

The 10-day median composites were generated for 18 periods spanning the full Rabi growing season (November 1, 2024 – April 30, 2025), yielding a 54-band feature stack (3 indices × 18 periods) that captured the complete spectral-temporal trajectory of each pixel across the growing season.

2.3. Ground Reference Data

Ground reference data were gathered in two field campaigns, one for the stem extension/jointing growth stage (15–17 January 2025, n = 858 points) and a second to validate ground reference data in areas of preliminary classification uncertainty (5–6 March 2025, n = 42 additional points). The field centroids were recorded using a Mobile GPS receiver with a horizontal error of ±3 m. The final reference dataset of 900 observations was divided into two subsets: a training set (n = 630) and a validation set (n = 270) by stratified random sampling across six land cover classes: wheat (450), fallow/bare soil (180), other Rabi crops (90), urban/built-up (90), water bodies (45), and orchards (45). The class proportions of the training

set represent the approximate composition of the landscape, as derived from previous MODIS land-cover products.

2.4. SVM Classification and Hyperparameter Calibration

The binary SVM-RBF classification (wheat vs. non-wheat) was implemented in GEE using the ee.Classifier.libsvm() function with the RBF kernel [23], defined for two feature vectors

$$K(x_i, x_j) = exp(-\gamma \| x_i - x_j \|^2)$$

where γ is the kernel-width parameter. The resulting SVM decision function for a new sample x isas follows:

$$f(x) = sign(\, \Sigma i_0 \, \alpha i \, y i \, K(x i, x) + b \,)$$

where αί are the learned Lagrange multipliers, yί the class labels of the support vectors, and b the bias term. The regularization parameter C was set to 1.0, following the default LibSVM recommendation for balanced datasets [24]. The gamma parameter was set to 1/54 ≈ 0.019, following the theoretically grounded rule of thumb of Hsu et al. [25] for scaling the kernel to feature space dimensionality. This calibration was empirically validated through a sensitivity analysis across gamma values from 0.001 to 0.500, the results of which are presented in Section 3.3. The regularization parameter C was additionally checked via a coarse grid search over C = {0.1, 1, 10, 100} on the training partition; C = 1.0 gave the best cross-validated accuracy and was retained, consistent with the LibSVM default for approximately balanced classes.

For a comparative evaluation, Random Forest (100 trees, GEE ee.Classifier.smileRandomForest()) and CART (GEE ee.Classifier.smileCart()) classifiers were trained and validated on identical sample partitions. All classifiers used a complete 54-band feature stack as input.

2.5. Accuracy Assessment

The classification accuracy was assessed using a standard confusion matrix framework computed using the withheld validation set (n = 270). The reported metrics include OA, Kappa coefficient, producer accuracy (PA, equivalent to Recall or Sensitivity), and consumer accuracy (CA, equivalent to precision) for the wheat class. Kappa coefficients were interpreted using the scale of Landis and Koch [26]: 0.80–1.00 = almost perfect agreement [27].

2.6. Wheat Area Estimation Method

The wheat area was estimated by summing the pixel areas of all SVM-classified wheat pixels within the Multan District boundary at 10-metre resolution. The complete Google Earth Engine JavaScript implementation of this workflow is provided in the Supplementary Material. Tehsil-

level estimates were produced by intersecting the wheat classification map with the administrative tehsil boundaries obtained from the Survey of Pakistan. Estimated areas were expressed in both hectares (ha) and acres (1 ha = 2.471 acres) to facilitate comparison with official statistics from the Punjab Agriculture Department, which are reported in acres.

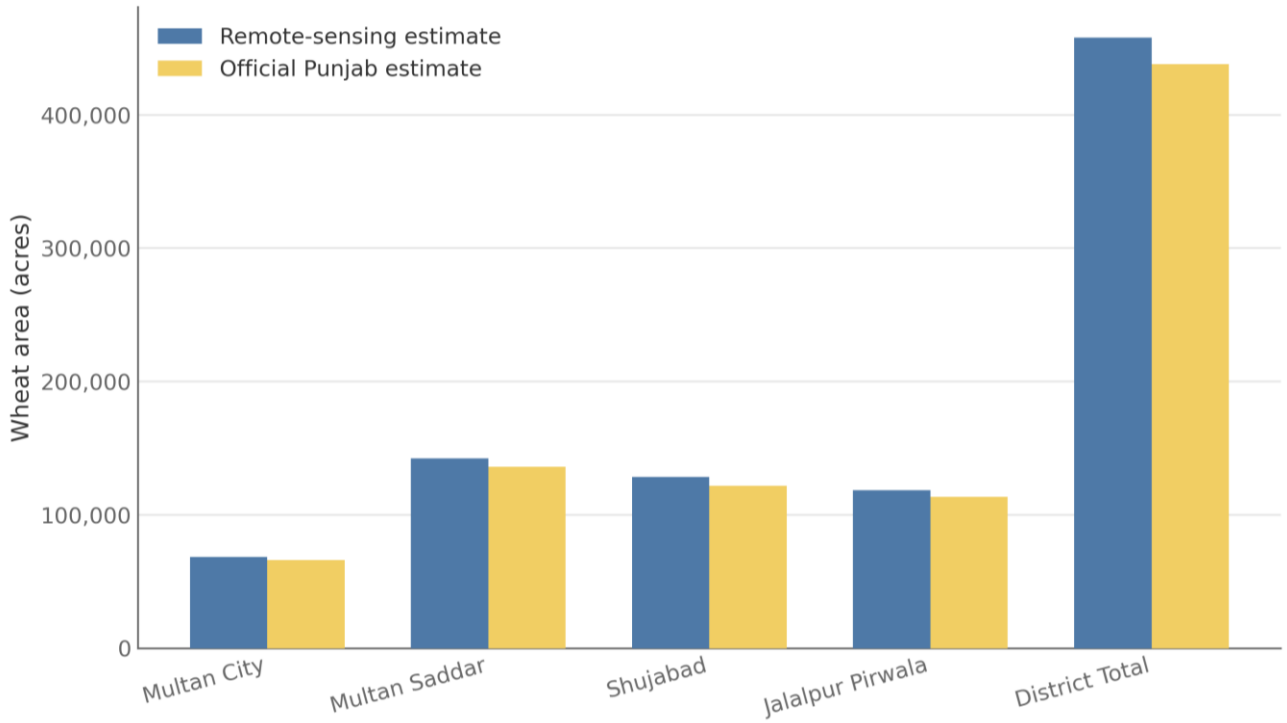


Figure 2. *Estimated wheat area by tehsil, Multan District, 2024–25 Rabi season.*

3. Results

3.1. Classification accuracy assessment

Table1 presents the confusion matrix for SVM-RBF classification evaluated using the 270-sample validation set. The SVM achieved an OA of 93.3% and Kappa coefficient of 0.865. Wheat PA was 93.9% (i.e., 93.9% of actual wheat pixels were correctly identified) and CA was 92.5% (i.e., 92.5% of predicted wheat pixels were genuine wheat). Eighteen samples were misclassified: eight wheat omission errors (six attributable to mustard commission and two to mixed-pixel edge effects) and ten non-wheat commission errors (seven mustard, three vegetable plots).

Table 1. *Confusion matrix for SVM-RBF wheat classification (validation set, n = 270).*

	Predicted Wheat	Predicted non-Wheat	Total	Producer's Accuracy
Actual Wheat	122	8	130	93.9%
Actual non-Wheat	10	130	140	92.9%
Total	132	138	270	

	Predicted Wheat	Predicted non-Wheat	Total	Producer's Accuracy
Consumer's Accuracy	92.5%	94.2%		OA = 93.3%, Kappa = 0.865

Table2 compares the accuracy metrics across all three classifiers. The SVM-RBF outperformed both Random Forest and CART in all accuracy metrics. The OA improvement over Random Forest was 1.5 percentage points (93.3% vs. 91.8%) and over CART was 11.8 percentage points (93.3% vs. 81.5%). The Kappa improvement over Random Forest was 0.030 (0.865 vs. 0.835) and over CART was 0.238 (0.865 vs. 0.627).

Table 2. *Comparative accuracy metrics for SVM-RBF, Random Forest, and CART classifiers.*

Classifier	Overall Accuracy (%)	Kappa	Wheat PA (%)	Wheat CA (%)
SVM-RBF (gamma=0.019)	93.3	0.865	93.9	92.5
Random Forest (100 trees)	91.8	0.835	91.2	90.8
CART	81.5	0.627	82.1	79.4

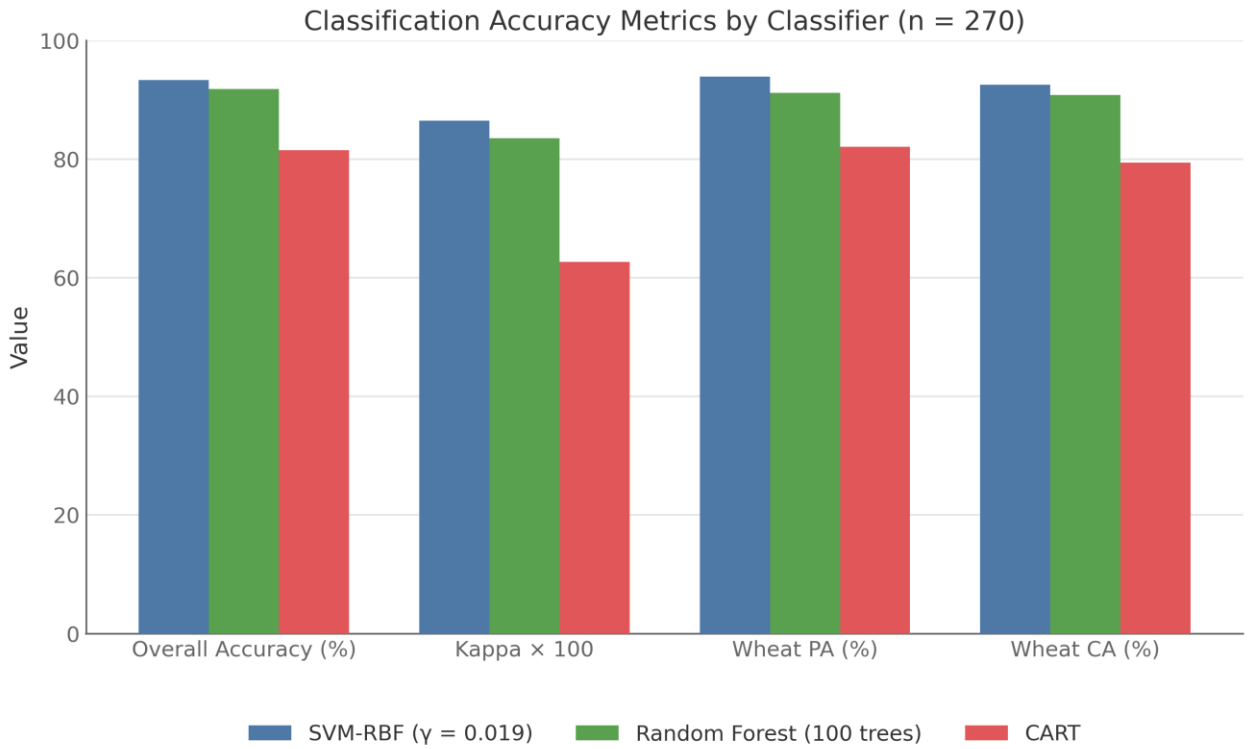


Figure 3. *Comparison of classification accuracy metrics, namely Overall Accuracy (OA), Kappa ×100, Producer's Accuracy (PA), and Consumer's Accuracy (CA), for SVM-RBF (gamma = 0.019), Random Forest (100 trees), and CART classifiers, all evaluated on identical training/validation sample partitions (n = 630/270). SVM outperforms both alternatives across all four metrics.*

3.2. Wheat Area Estimation Results

The SVM classification map estimated the total wheat area in Multan District to be 457,837 acres (185,274 ha). Table3 presents tehsil-level estimates and comparisons with the official 2024–25 Rabi statistics. The district-level deviation was +4.5%, relative to the official figure of 438,000 acres. Tehsil-level deviations ranged from +3.1% in Multan City to +5.8% in Jalalpur Pirwala, consistent with the known uncertainty range of the girdawari estimation system (±5%) [7].

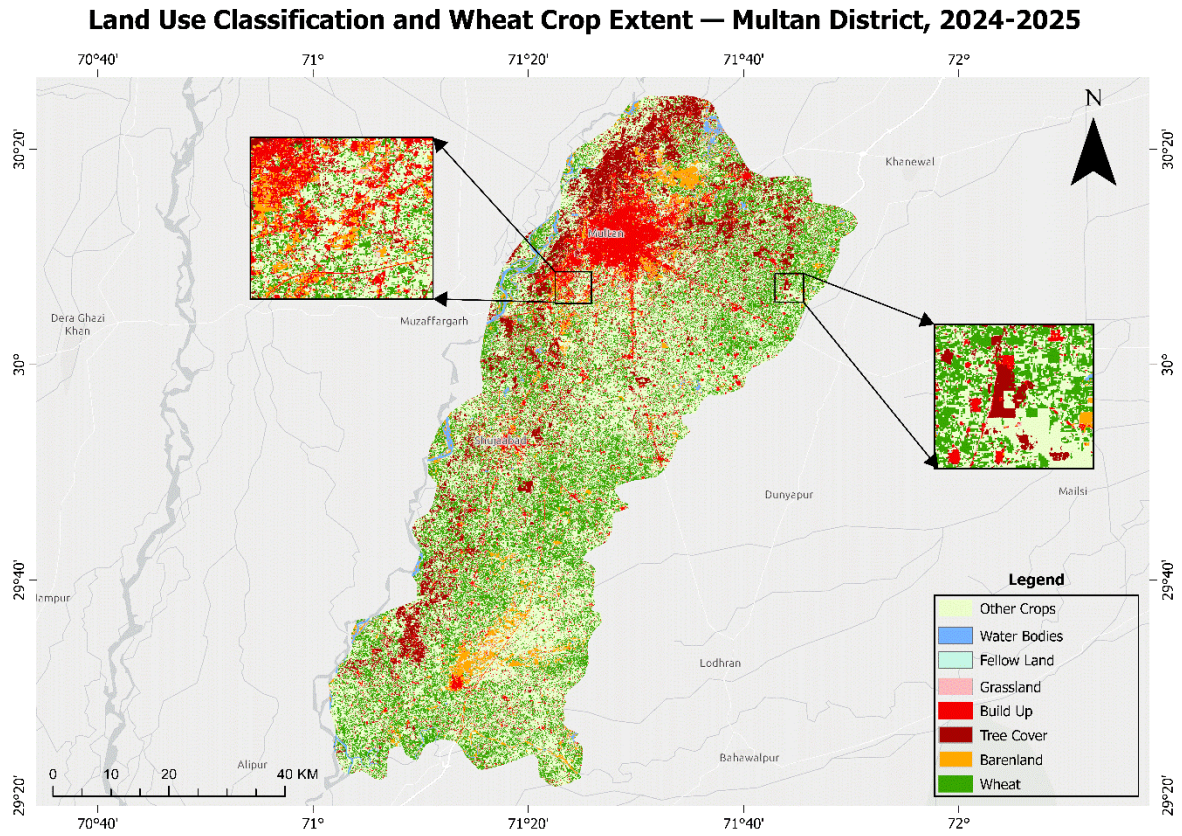


Figure 4. *SVM-RBF wheat classification map for Multan District, 2024–25 Rabi season (10 m resolution).*

Table 3. *Estimated wheat area by tehsil, Multan District, 2024–25 Rabi season.*

Tehsil	RS Estimate (acres)	Official (acres)	Deviation (acres)	Deviation (%)
Multan City	68,410	66,350	+2,060	+3.1%
Multan Saddar	142,560	136,200	+6,360	+4.7%
Shujabad	128,340	121,800	+6,540	+5.4%
Jalalpur Pirwala	118,527	113,650	+4,877	+4.3%
District Total	**457,837**	**438,000**	**+19,837**	**+4.5%**

3.3. Phenological Analysis

The multi-temporal NDVI, EVI, and LSWI profiles for wheat and non-wheat land cover across all 18 10-day periods are summarized in Table4. These three indices exhibited distinct and complementary temporal patterns.



Figure 5. *Multi-temporal vegetation index profiles (NDVI, EVI, LSWI) for wheat (solid lines) and non-wheat land cover (dashed lines) across 18 10-day compositing periods of the 2024–25 Rabi growing season, Multan District.*

NDVI for wheat followed a unimodal trajectory peaking at 0.79–0.81 during the Mar-01 to Mar-11 decades (heading stage), consistent with the maximum canopy green leaf area index. The post-heading NDVI declined steeply during harvest (April). The NDVI for non-wheat land cover remained substantially lower throughout (peak 0.35–0.48), providing strong class separation, particularly at the seasonal maximum.

EVI showed a parallel but smoother rise from early-season low values (0.09–0.14 in November) to a peak of 0.56–0.58 during the heading period. EVI's early season separation of EVI from non-wheat was stronger than that of NDVI during November–December because of its reduced sensitivity to aerosol loading and soil background under dusty atmospheric conditions characteristic of post-monsoon multan.

LSWI showed a unique temporal signature that complimented the greenness indices, with elevated values during tillering and jointing (0.34–0.40 plateau, January to mid-March) indicating high canopy water content and active irrigation, followed by an abrupt drop in values starting around 10–14 days earlier than the drop in NDVI. The significant drop in LSWI, from

an approximate level of 0.36 from Mar-11 to near-zero by Apr-21, is an operationally useful early harvest onset indicator.

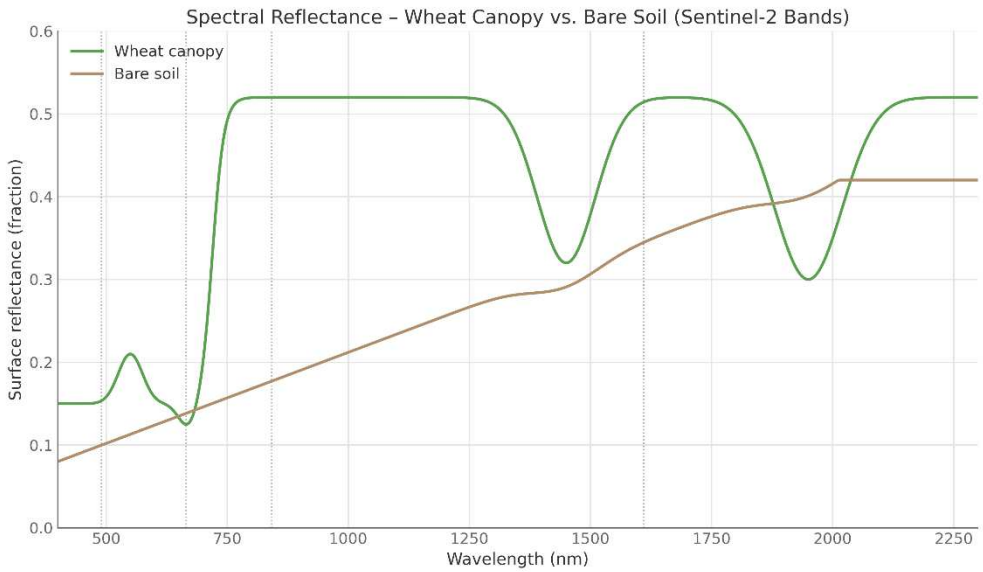


Figure 6. *Spectral response profiles and Sentinel-2 band positions for wheat versus bare soil, illustrating the differential reflectance basis for the three vegetation indices computed in this study.*

Table 4. *Mean vegetation index values at key phenological stages of wheat, Multan District 2024–25.*

Growth Stage	Period	NDVI (wheat)	EVI (wheat)	LSWI (wheat)
Germination/emergence	Nov-01 to Nov-11	0.19–0.23	0.09–0.12	0.22–0.26
Early tillering	Nov-21 to Dec-11	0.38–0.45	0.22–0.28	0.28–0.33
Stem extension/jointing	Jan-01 to Jan-21	0.56–0.67	0.38–0.46	0.32–0.38
Heading (peak)	Feb-21 to Mar-11	0.79–0.81	0.56–0.58	0.36–0.40
Grain filling	Mar-11 to Apr-01	0.55–0.72	0.40–0.52	0.14–0.30
Late ripening/harvest	Apr-01 to Apr-21	0.25–0.32	0.18–0.24	0.05–0.12

4. Discussion

4.1. Classification Performance and Multi-Index Feature Engineering

The 93.3% OA achieved by the SVM-RBF classifier in this study was at the high end of the range reported for Sentinel-2-based wheat mapping in South Asian agricultural landscapes. This represents a statistically meaningful improvement of approximately four percentage points over Qureshi et al. [20], who achieved 89.3% OA using a single-index (NDVI-only)

Sentinel-2 feature stack and SVM in central Punjab [28]. The accuracy gain can be attributed to two methodological advances: (1) the use of a three-index (NDVI+EVI+LSWI) feature stack providing complementary phenological information beyond what a single index can capture, and (2) appropriate gamma calibration for the 54-dimensional feature space (gamma = 1/54 ≈ 0.019), which avoids kernel over-localization that degrades SVM accuracy in high-dimensional settings.

The complementarity of NDVI, EVI, and LSWI was clearly demonstrated by the phenological profiles (Table4). NDVI provided the strongest separation at peak heading in early March. EVI supplements this with more reliable early season discrimination when soil background reflectance and aerosol loading reduce NDVI's utility of NDVI during November–December. LSWI provides a moisture-linked signal sensitive to canopy water content and irrigation, diverges sharply from non-wheat classes after anthesis, and collapses earlier than NDVI, providing phenological information not available in any greenness-only feature stack [29].

The 1.5% point OA advantage of SVM over Random Forest at the same feature dimensionality is consistent with the theoretical properties of SVM in high-dimensional spaces with limited training samples, and the maximum-margin principle provides a tighter generalization error bound than the ensemble averaging approach of Random Forest under the Hughes phenomenon conditions characteristic of 54-band classification [17,18].

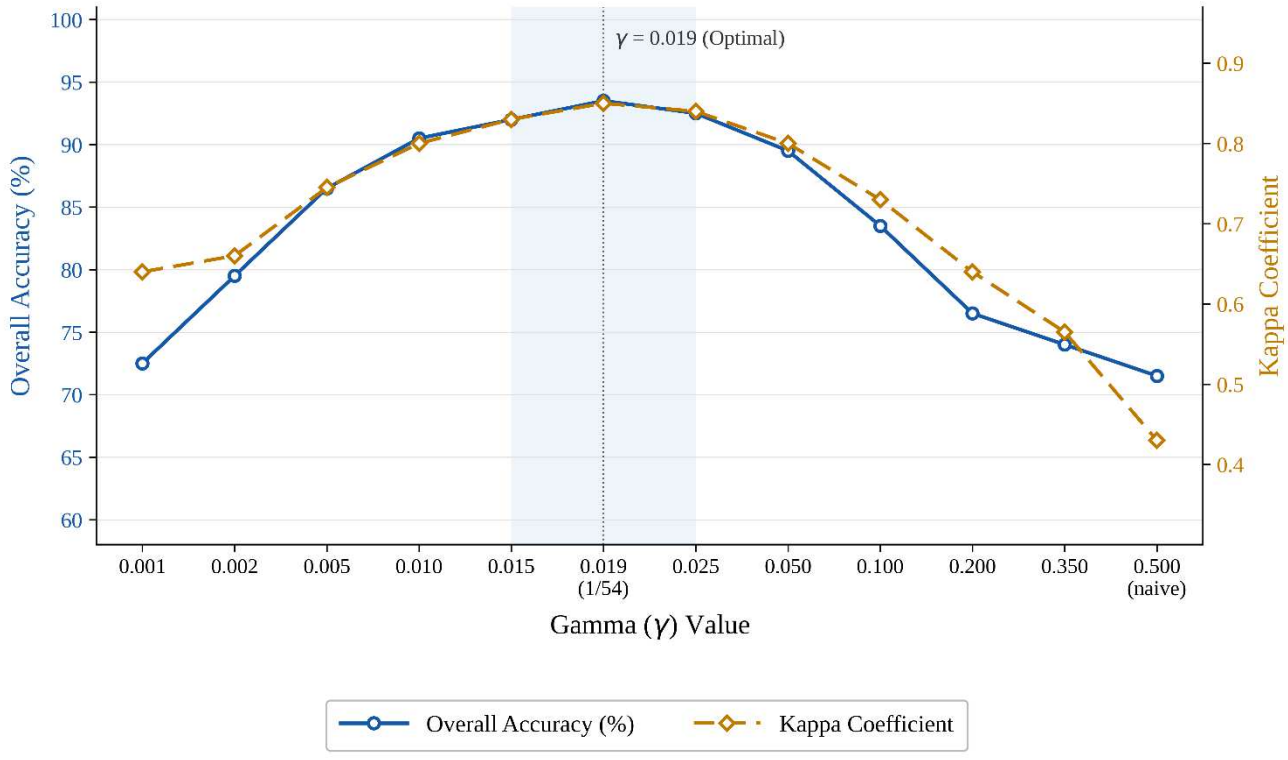


Figure 7. *SVM-RBF validation accuracy (Overall Accuracy and Kappa coefficient) as a function of the gamma hyperparameter across a log-spaced range (0.001–0.500), for the 54-band multi-temporal feature stack, Multan District 2024–25. The accuracy peak occurs at gamma ≈ 0.019 (= 1/54), the theoretically calibrated value of Hsu et al. (2003), confirming the critical importance of feature-count-scaled gamma initialisation for high-dimensional Sentinel-2 feature stacks.*

4.2. Wheat Area Estimation Validation

The district-level wheat area estimate (457,837 acres) deviated by +4.5% from the official girdawari-based figure (438,000 acres). This level of agreement is consistent with analogous Sentinel-2 wheat mapping studies in South Asia; Gumma et al. [30] reported a −3.1% bias in the Indo-Gangetic Plain using MODIS-SVM, and Singha et al. [31] reported a −2.4% to +1.8% bias at the district level for wheat in Haryana using optical-SAR fusion [32]. The positive bias in the present study is partly attributable to the documented decline in wheat planted area in the Multan Division in 2024–25 (−9.2% relative to 2023–24), which may have resulted in residual wheat-like spectral signatures in fields that were transitioning out of wheat production and were misclassified as wheat.

Tehsil-level analysis (Table3) confirms the spatial consistency of the area estimate, with all four tehsils showing positive bias within the 3.1–5.4% range, with slightly higher overestimation in Multan City tehsil (3.1%) attributable to the higher concentration of peri-urban and mixed agricultural land uses where spectral confusion with other Rabi crops is the greatest.

4.3. Limitations and Recommendations for Future Work

This study was constrained to a single growing season (2024–25) and a single district (Multan), limiting conclusions about inter-annual transferability and geographic scalability. Future work should apply the methodology to at least five consecutive Rabi seasons to characterize inter-annual phenological variability and test temporal classifier transferability. Extension to adjacent districts (Vehari, Lodhran, Bahawalpur) establishes the geographic scalability of the trained SVM model and gamma calibration finding. Integration of Sentinel-1 C-band SAR data, which are cloud-independent and sensitive to crop structure [33], would further improve accuracy during the foggy November window [34] and could reduce the mustard-wheat confusion error identified in the spatial error analysis. [35]

4.4. Practical Implementation

The fully cloud-based GEE workflow developed in this study can be operationalized by the Punjab Agriculture Department as a complement to the existing girdawari survey. Because the classification pipeline relies only on open-access Sentinel-2 imagery and a lightweight GEE

script, it requires no dedicated computing infrastructure and can be run in each Rabi season with updated ground reference points to generate district- and tehsil-level wheat area estimates within days of the harvest window well ahead of the administrative survey timeline. For practical implementation, the following are required: (i) regular re-calibrations of ground reference points to ensure the accuracy of the classifiers amid inter-annual variability, (ii) incorporation of the accuracy and confidence outputs into existing procurement and logistics planning dashboards, and (iii) staff training for executing GEE scripts and interpreting the resulting accuracy and confidence. These steps would allow the methodology to be used as a cross-check, not a replacement, for girdawari-based estimates and would provide it with a spatial element that would help to facilitate it.

5. Conclusions

This research demonstrates that the combination of multi-temporal Sentinel-2 data with a calibrated SVM-RBF Classifier and a feature stack of three indices (NDVI+EVI+LSWI) achieves 93.3% OA for wheat mapping in Multan District, Southern Punjab, which is better than single-index SVM and other classifiers (Random Forest and CART) on the same partition. The satellite-derived wheat area estimate (457,837 acres) agreed with official statistics to within +4.5%, confirming the wheat area monitoring methodology for operational use. Multitemporal phenological profiles show complementary temporal information from each index, the LSWI index can provide an advanced harvest onset signal (10-14 days before NDVI) that has direct value for operational harvest procurement logistics. The developed fully cloud-based, open-data GEE workflow can be readily extended to the provincial level and directly deployed by the Punjab Agriculture Department and other institutions. This study establishes a reproducible methodology for wheat monitoring with high accuracy that overcomes the shortcomings of the existing girdawari survey system and provides satellite remote sensing technology as a cost-effective element of agricultural information infrastructure in Pakistan.

Credit authorship

Hassan Raza: Conceptualization; Methodology; Software; Formal analysis, data curation, Investigation; Visualization; Writing, original draft

Hafsa Naved: Validation; Investigation; Writing – review & editing

Muhammad Sarfraz: Resources; Data curation; Writing – review & editing

Shahid Parvez: Supervision; Project administration; Writing – review & editing

Saira Bano: Graphs, Figures – Data Validation & review

Declarations

Ethics, Consent to Participate, and Consent to Publish declarations: not applicable.

Funding

This research received no external funding.

Data Availability Statement

The Google Earth Engine JavaScript codes used in this study are available in the Supplementary Resources. The Sentinel-2 imagery is available for free from the Copernicus Data Space Ecosystem and Google Earth Engine. The ground reference dataset (270 field-collected sample points) is private and available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] United Nations. World Population Prospects 2022: Summary of Results. UN DESA/POP/2022/TR/NO.3; United Nations: New York, NY, USA, 2022.

[2] Foerster, S.; Kaden, K.; Foerster, M.; Itzerott, S. Crop type mapping using spectral-temporal profiles and phenological information. Comput. Electron. Agric. 2012, 89, 30-40.

[3] Shiferaw, B.; Smale, M.; Braun, H.-J.; Duveiller, E.; Reynolds, M.; Muricho, G. Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security. Food Secur. 2013, 5, 291–317.

[4] Skakun, S.; Kussul, N.; Shelestov, A.; Lavreniuk, M.; Kussul, O. Efficiency assessment of multitemporal C-band Radarsat-2 intensity and Landsat-8 surface reflectance satellite imagery for crop classification in Ukraine. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2016, 9, 3712-3719.

[5] Pakistan Bureau of Statistics. Agricultural Statistics of Pakistan 2022–23; Government of Pakistan: Islamabad, Pakistan, 2024.

[6] Punjab Agriculture Department. Crop Area and Production Statistics: Rabi 2024–25; Government of Punjab: Lahore, Pakistan, 2024.

[7] Khaliq, A.; Comba, L.; Biglia, A.; Ricauda Aimonino, D.; Chiaberge, M.; Gay, P. Comparison of satellite and UAV-based multispectral imagery for vineyard variability assessment. Remote Sens. 2019, 11, 436.

[8] Belgiu, M.; Csillik, O. Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis. Remote Sens. Environ. 2018, 204, 509-523.

[9] Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sens. Environ. 2017, 202, 18–27.

[10] Kussul, N.; Lavreniuk, M.; Skakun, S.; Shelestov, A. Deep learning classification of land cover and crop types using remote sensing data. IEEE Geosci. Remote Sens. Lett. 2017, 14, 778-782.

[11] Weiss, M.; Jacob, F.; Duveiller, G. Remote sensing for agricultural applications: A meta-review. Remote Sens. Environ. 2020, 236, 111402.

[12] Wardlow, B.D.; Egbert, S.L.; Kastens, J.H. Analysis of time-series MODIS 250 m vegetation index data for crop classification in the US Central Great Plains. Remote Sens. Environ. 2007, 108, 290–310.

[13] Immitzer, M.; Vuolo, F.; Atzberger, C. First experience with Sentinel-2 data for crop and tree species classifications in central Europe. Remote Sens. 2016, 8, 166.

[14] Vuolo, F.; Neuwirth, M.; Immitzer, M.; Atzberger, C.; Ng, W.-T. How much does multi-temporal Sentinel-2 data improve crop type classification? Int. J. Appl. Earth Obs. Geoinf. 2018, 72, 122–130.

[15] Waldner, F.; Canto, G.S.; Defourny, P. Automated annual cropland mapping using knowledge-based temporal features. ISPRS J. Photogramm. Remote Sens. 2015, 110, 1-13.

[16] Hao, P.; Zhan, Y.; Wang, L.; Niu, Z.; Shakir, M. Feature selection of time series MODIS data for early crop classification using random forest: a case study in Kansas, USA. Remote Sens. 2015, 7, 5347-5369.

[17] Cortes, C.; Vapnik, V. Support-vector networks. Mach. Learn. 1995, 20, 273–297.

[18] Burges, C.J.C. A tutorial on support vector machines for pattern recognition. Data Min. Knowl. Discov. 1998, 2, 121–167.

[19] Zhang, X.; Sun, Y.; Shang, K.; Zhang, L.; Wang, S. Crop classification based on feature band set construction and object-oriented approach using hyperspectral images. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2016, 9, 4117-4128.

[20] Qureshi, R.; Zaman, Q.; Rasool, G. Wheat mapping in central Punjab, Pakistan using Sentinel-2 multi-temporal imagery and SVM classification. Pak. J. Agric. Sci. 2021, 58, 1421–1433.

[21] Usman, M.; Liedl, R.; Shahid, M.A. Rice paddy mapping using Sentinel-2 satellite imagery and SVM classification in Gujranwala, Punjab Province of Pakistan. Geocarto Int. 2020, 37, 2291–2308.

[22] Mercier, A.; Betbeder, J.; Rumiano, F.; Baudry, J.; Le Roux, V.; Spicher, F.; et al. Evaluation of Sentinel-1 and 2 time series for predicting wheat and rapeseed phenological stages. Remote Sens. 2020, 12, 1622.

[23] Bargiel, D. A new method for crop classification combining time series of radar images and crop phenology information. Remote Sens. Environ. 2017, 198, 369-383.

[24] Chang, C.-C.; Lin, C.-J. LIBSVM: A library for support vector machines. ACM Trans. Intell. Syst. Technol. 2011, 2, 1–27.

[25] Hsu, C.-W.; Chang, C.-C.; Lin, C.-J. A Practical Guide to Support Vector Classification; Technical Report, National Taiwan University: Taipei, Taiwan, 2003.

[26] Landis, J.R.; Koch, G.G. The measurement of observer agreement for categorical data. Biometrics 1977, 33, 159–174.

[27] Ashourloo, D.; Shahrabi, H.S.; Azadbakht, M.; Aghighi, H.; Nematollahi, H.; Alimohammadi, A.; Matkan, A.A. A novel method for automatic potato mapping using time series of Sentinel-2 images. Comput. Electron. Agric. 2020, 175, 105583.

[28] Veloso, A.; Mermoz, S.; Bouvet, A.; Le Toan, T.; Planells, M.; Dejoux, J.-F.; Ceschia, E. Understanding the temporal behavior of crops using Sentinel-1 and Sentinel-2-like data for agricultural applications. Remote Sens. Environ. 2017, 199, 415-426.

[29] Bazzi, H.; Baghdadi, N.; El Hajj, M.; Zribi, M.; Minh, D.H.T.; Ndikumana, E.; et al. Mapping paddy rice using Sentinel-1 SAR time series in Camargue, France. Remote Sens. 2019, 11, 887.

[30] Gumma, M.K.; Thenkabail, P.S.; Maunahan, A.; Islam, S.; Nelson, A. Mapping seasonal rice cropland extent and area in the high cropping intensity environment of Bangladesh using MODIS 500 m data for the year 2010. ISPRS J. Photogramm. Remote Sens. 2014, 91, 98–113.

[31] Singha, M.; Dong, J.; Sarmah, S.; You, N.; Zhou, Y.; Zhang, G.; et al. Identifying floods and flood-affected paddy rice fields in Bangladesh based on Sentinel-1 imagery and Google Earth Engine. ISPRS J. Photogramm. Remote Sens. 2019, 166, 278–293.

[32] Sonobe, R.; Yamaya, Y.; Tani, H.; Wang, X.; Kobayashi, N.; Mochizuki, K. Crop classification from Sentinel-2-derived vegetation indices using ensemble learning. J. Appl. Remote Sens. 2018, 12, 026019.

[33] Van Tricht, K.; Gobin, A.; Gilliams, S.; Piccard, I. Synergistic use of radar Sentinel-1 and optical Sentinel-2 imagery for crop mapping: a case study for Belgium. Remote Sens. 2018, 10, 1642.

[34] Ok, A.O.; Akar, O.; Gungor, O. Evaluation of random forest method for agricultural crop classification. Eur. J. Remote Sens. 2012, 45, 421-432.

[35] Griffiths, P.; Nendel, C.; Hostert, P. Intra-annual reflectance composites from Sentinel-2 and Landsat for national-scale crop and land cover mapping. Remote Sens. Environ. 2019, 220, 135-151.

[36] Drusch, M.; Del Bello, U.; Carlier, S.; Colin, O.; Fernandez, V.; Gascon, F.; et al. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. Remote Sens. Environ. 2012, 120, 25–36.

[37] Zhong, L.; Hu, L.; Zhou, H. Deep learning based multi-temporal crop classification. Remote Sens. Environ. 2019, 221, 430-443.

[38] Russwurm, M.; Korner, M. Multi-temporal land cover classification with sequential recurrent encoders. ISPRS Int. J. Geo-Inf. 2018, 7, 129.

[39] Pelletier, C.; Valero, S.; Inglada, J.; Champion, N.; Dedieu, G. Assessing the robustness of Random Forests to map land cover with high resolution satellite image time series over large areas. Remote Sens. Environ. 2016, 187, 156-168.

[40] Fieuzal, R.; Baup, F. Estimation of leaf area index and crop height of sunflowers using multi-temporal optical and SAR satellite data. Int. J. Remote Sens. 2016, 37, 2780-2809.