Predicting Rice Yield Using Multi-Temporal Vegetation Indices and Machine Learning: A Comparative Study of Random Forest and Support Vector Regression Models

Rabia Shaheen[1], Zeeshan Khan[1]
[1]Department of Computer Science, Superior University, Lahore
*Correspondence: Shaheen.r@gmail.com

Citation | Shaheen. R, Khan. Z, "Predicting Rice Yield Using Multi-Temporal Vegetation Indices and Machine Learning: A Comparative Study of Random Forest and Support Vector Regression Models", FCSI, Vol. 02 Issue. 4 pp 158-167, Oct 2024
Received | Sep 09, 2024, **Revised |** Oct 10, 2024, **Accepted |** Oct 12, 2024, **Published |** Oct 13, 2024.

Accurate prediction of rice yield is essential for ensuring food security and optimizing agricultural management practices. This study integrates multi-temporal remote sensing vegetation indices—Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), and Normalized Difference Red Edge (NDRE)—with advanced machine learning algorithms to forecast rice yield. Time-series data from critical phenological stages were analyzed to capture crop growth dynamics, while Random Forest (RF) and Support Vector Regression (SVR) models were developed to estimate yield outcomes. Results indicate that the RF model outperforms SVR in predictive accuracy, with an R^2 value of 0.92 compared to 0.87 for SVR. Feature importance analysis identified maximum NDVI and cumulative NDVI (area under curve) as the most influential predictors, emphasizing the significance of canopy vigor and growth duration. Spatial yield maps generated from the RF model provide valuable insights for precision agriculture interventions. The proposed framework demonstrates robust potential for early-season yield forecasting, facilitating improved resource allocation and decision-making in rice cultivation. Future research should focus on expanding the model's applicability across different environments and integrating real-time sensor data for dynamic crop monitoring.

Keywords: Rice Yield Prediction, Remote Sensing, Vegetation Indices, NDVI, EVI, NDRE, Machine Learning, Random Forest, Support Vector Regression, Time-Series Data

Introduction:

Precision agriculture (PA) has emerged as a transformative approach to meeting the growing global demand for food while conserving resources and reducing environmental impacts. By leveraging advanced technologies such as remote sensing (RS), geospatial analytics, and artificial intelligence (AI), PA enables farmers to make site-specific management decisions that optimize yields and resource efficiency. Among RS tools, the Sentinel-2 satellite system—developed under the European Space Agency's Copernicus program—offers unique advantages due to its high spatial resolution (10–20 m), wide spectral range across 13 bands, and frequent revisit period of five days. These capabilities facilitate detailed monitoring of crop phenology, health, and stress, supporting timely agricultural interventions and yield forecasting.

Vegetation indices (VIs) such as the Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), and Normalized Difference Red Edge Index (NDRE) derived from Sentinel-2 imagery have become indispensable in crop monitoring and yield estimation. When integrated with AI models—particularly machine learning (ML) and deep learning (DL) algorithms—these datasets can capture complex, non-linear relationships between environmental variables and yield outcomes, enabling accurate predictions at both field and regional scales. This fusion of Sentinel-2 data and AI is particularly valuable for diverse cropping systems, from staple cereals like wheat and rice to specialized crops such as potatoes and rubber, where yield variability is influenced by climate, soil, and management practices.

Despite growing adoption, variations exist in the vegetation indices used, the AI algorithms applied, and the temporal strategies for monitoring (seasonal snapshots vs. continuous observation). These methodological differences can significantly influence prediction accuracy, creating the need for comparative assessments and methodological standardization in Sentinel-2–based yield prediction research.

Research Gap:

Although several review articles have examined AI-based yield prediction in agriculture, none have specifically focused on Sentinel-2 applications despite its increasing prominence in precision agriculture over the past five years. Existing studies often treat Sentinel-2 as one of many RS sources, providing only limited discussion of its unique spectral, spatial, and temporal advantages. Moreover, there is no consolidated analysis of which vegetation indices, AI/ML models, and crop types yield the most accurate results when using Sentinel-2 data. While recent works have explored the combination of Sentinel-2 with other platforms (e.g., Landsat 8, UAV imagery), there is insufficient synthesis on when and why multisource integration is beneficial versus when Sentinel-2 alone suffices. Furthermore, methodological inconsistencies—such as differences in feature selection, phenological stage targeting, and model calibration—hinder cross-study comparability and limit the development of generalizable best practices for Sentinel-2–based yield forecasting.

Objectives:

This study aims to provide a comprehensive review of research conducted between 2019 and 2024 on crop yield prediction using Sentinel-2 satellite data, with a particular focus on the integration of vegetation indices and artificial intelligence (AI) techniques. The primary objectives include analyzing recent literature to identify the most commonly employed vegetation indices, AI and machine learning (ML) models, and temporal monitoring strategies utilized in Sentinel-2–based yield prediction studies. Additionally, the study compares methodological variations across different research works and evaluates how these differences impact prediction accuracy for various crop types and geographic regions. Another key focus is assessing the role of Sentinel-2 data in single-source approaches as well as in multisource data integration methods for agricultural yield forecasting.

Novelty Statement:

This is the first review study to focus exclusively on Sentinel-2–based crop yield prediction, covering the most recent advances between 2019 and 2024. Unlike broader reviews of AI in agriculture, this work systematically evaluates the spectral, spatial, and temporal strengths of Sentinel-2, identifies optimal vegetation indices and AI models for specific crops, and examines the methodological trade-offs between continuous monitoring and seasonal analysis. By synthesizing findings across diverse agro-ecological zones, the study provides a framework for selecting appropriate Sentinel-2–AI combinations tailored to crop type, growth stage, and resource constraints. This targeted approach will bridge the current gap between technological capability and practical adoption in field-level decision-making, contributing to improved yield forecasting, resource optimization, and sustainable agricultural practices.

Recent literature underscores the critical role of Sentinel-2 in enabling accurate, high-resolution yield forecasts when combined with AI [1][2][3][4][5]. These studies highlight the potential for Sentinel-2 to serve as a primary data source for global yield monitoring initiatives, reinforcing the need for a specialized, methodology-focused synthesis.

Literature review:

Recent advances in precision agriculture have increasingly leveraged **Sentinel-2** satellite imagery—notably its high spatial resolution (10–20 m), wide spectral range (13 bands including red-edge), and frequent revisit intervals—to improve crop monitoring and yield prediction [6][7][8]. Vegetation indices (VIs) derived from Sentinel-2 data—such as NDVI, EVI, and red-edge indices—have proven highly effective as proxies for vegetation health, biomass, and stress, serving as foundational features for both traditional machine learning (ML) and modern deep learning (DL) models [7][9].

Two methodological trends are notable. The first employs a simpler VI-focused ML approach, using selected indices at key phenological stages and feeding them into interpretable models like Random Forests (RF), Support Vector Regression (SVR), or Partial Least Squares Regression (PLSR). These methods offer robustness and lower data requirements; for instance, studies in Ethiopia demonstrated that NDVI and EVI effectively predicted teff and finger millet yield with $R^2 \approx 0.84$–0.87 and RMSE values below 1 t/ha [9]. The second trend involves more complex models: deep learning architectures, such as CNNs, RNNs, and attention-based networks, trained on Sentinel-2 time series. These architectures can capture temporal and spatial patterns, often outperforming ML when sufficient labeled data is available [7][8].

Integrative approaches that combine Sentinel-2 data with ancillary datasets—including soil variables, weather/climate reanalysis, UAV imagery, or SAR data—are gaining prominence. Such fusion enhances prediction accuracy and early forecasting capability. For example, combining optical and SAR data improved yield estimates across developmental stages beyond what VI maxima alone could provide [10]. Similarly, integrating Sentinel-2 with time-aligned UAV imagery has enhanced model performance in complex terrains [11].

Yet, key challenges persist. Methodological inconsistencies—such as varying VI selections, phenological timing, preprocessing routines (e.g., cloud masking, gap filling), and ground-truth sampling—limit cross-study comparability and model transferability. Moreover, while DL offers higher performance, its data and computational demands restrict adoption in regions with limited labeled data [7][8]. To address these gaps, recent literature advocates for standardized benchmarking datasets, transparent documentation, and systematic comparisons between single-source Sentinel-2 workflows and multisource fusion approaches across diverse crops and agro-ecologies [7].

In summary, Sentinel-2 imagery—when combined with well-selected VIs, AI models, and complementary data—holds considerable promise for scalable, accurate yield prediction. Yet, realizing this potential requires methodological rigor, reproducibility, and adaptability across crops and environments.

Methodology:

Study Area:

The study was conducted in the Lambayeque region, Peru, located between latitudes 6°24'S and 7°01'S and longitudes 79°49'W and 80°22'W. The region exhibits a semi-arid climate with average annual precipitation of approximately 150–200 mm, primarily occurring between December and March. Rice (*Oryza sativa L.*) is the dominant crop, supported by extensive irrigation networks fed by the Chancay-Lambayeque River. The selection of this site was based on the presence of spatial variability in planting dates, irrigation schedules, and soil characteristics, providing a robust setting for testing remote sensing–based yield estimation.

Data Sources and Acquisition:

Data Collection:

Sentinel-2 Level-2A satellite imagery for the 2022–2023 growing season (October 2022 to April 2023) was obtained from the Copernicus Open Access Hub. The imagery featured a spatial resolution of 10 meters for visible bands (B2, B3, B4) and near-infrared band (B8), and

20 meters for red-edge (B5, B6, B7, B8A) and shortwave infrared bands (B11, B12). A temporal resolution of five days was maintained, and images with less than 10% cloud cover, as indicated by the Scene Classification Layer (SCL), were selected for analysis. Ground truth yield data were collected from 36 representative rice plots ranging between 0.5 and 1 hectare. At physiological maturity, a 5 m × 5 m quadrat was harvested from each plot. Grain samples were weighed immediately in the field and then oven-dried at 70 °C to constant mass. The resulting dry weights were converted into tons per hectare (t/ha). All sampling locations were precisely georeferenced using a Garmin GPSMAP 64sx device with an accuracy of approximately ±3 meters.

Satellite Data Preprocessing:

Sentinel-2 Level-2A products underwent atmospheric correction using the Sen2Cor processor to generate bottom-of-atmosphere reflectance values. Cloud-affected pixels were masked using the SCL, with a 60-meter morphological buffer applied around cloud edges to minimize contamination. To ensure continuous phenological time series, temporal gaps caused by missing data were filled via linear interpolation between valid observations. For each acquisition date, five vegetation indices (VIs) were calculated: Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), Normalized Difference Red Edge (NDRE),

Soil Adjusted Vegetation Index (SAVI, with L = 0.5), and Green Chlorophyll Vegetation Index (GCVI). These indices were extracted from a 3 × 3 pixel window centered on the GPS coordinates of each ground-truth plot.

To accurately capture crop phenology, NDVI time series were smoothed using a Savitzky–Golay filter (window size of 5) to reduce noise. Key phenological stages—tillering, panicle initiation, flowering, and grain filling—were identified for each plot. At these stages, the corresponding VI values were recorded, along with seasonal metrics such as maximum value, mean value, and area under the curve (AUC), providing comprehensive spectral characterizations for model input.

Model Development and Evaluation

Feature selection was conducted using Pearson correlation analysis to identify spectral and vegetation index metrics strongly associated with observed yields. Features exhibiting high collinearity (correlation coefficient $|r| > 0.85$) were excluded to prevent multicollinearity issues. Two machine learning regression algorithms were employed to predict crop yields: Random Forest Regression (RF) and Support Vector Regression (SVR). RF was selected for its robustness to overfitting and capacity to model complex non-linear relationships, while SVR was applied to evaluate the performance of kernel-based regression methods in this context. Model hyperparameters were optimized through grid search combined with 10-fold cross-validation on the training dataset. The dataset was split into training (70%) and testing (30%) subsets to assess model generalization. Model performance metrics included the coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE).

Software and Analytical Tools:

Satellite data preprocessing and vegetation index calculations were performed using Google Earth Engine (GEE), while spatial visualization and coordinate matching were conducted in QGIS version 3.28. Statistical analysis, machine learning model training, and performance evaluation were implemented in Python 3.10, utilizing libraries such as scikit-learn, pandas, numpy, and matplotlib.

Results:

The field-based yield measurements collected from the 36 rice plots exhibited substantial variability, ranging from 4.21 t/ha in poorly irrigated fields to 8.94 t/ha in optimally managed ones, with a mean yield of 6.47 t/ha (SD = 1.21). Early-planted fields, particularly those established within the first two weeks of the planting window, demonstrated an average yield advantage of approximately 1.12 t/ha over late-planted fields [12]. ANOVA results indicated that planting date was a statistically significant determinant of yield ($F(1,34) = 8.62$, $p < 0.01$). Irrigation uniformity and fertilizer application frequency also showed a positive association with yield performance, while water stress during the panicle initiation stage was consistently linked to yield reductions of 15–22%.

Time-series analysis of Sentinel-2 spectral indices revealed well-defined phenological curves across the growing season. NDVI values gradually increased from a mean of 0.45 (±0.04) during the tillering stage to peak values averaging 0.89 (±0.02) at flowering, before declining to 0.65 at maturity as chlorophyll content decreased. EVI followed a similar trend, peaking at 0.73, but with slightly less sensitivity to full canopy closure [13]. NDRE showed an earlier peak at 0.58, reflecting its responsiveness to chlorophyll content before the canopy reached full density. Seasonal metrics such as NDVI_max, NDVI_AUC, and NDRE_mean exhibited strong positive correlations with observed yields ($r = 0.81$, $r = 0.76$, and $r = 0.74$, respectively; $p < 0.01$). SWIR-based indices such as SAVI and GCVI displayed moderate correlations (0.55–0.61), likely due to their mixed sensitivity to both biomass and soil moisture variations.

Multicollinearity diagnostics using Variance Inflation Factor (VIF) analysis refined the predictor set from 25 initial vegetation indices and spectral metrics to 11 key variables, ensuring that no retained predictor exceeded a VIF value of 5. Feature importance analysis from the Random Forest (RF) model identified NDVI_max (importance score = 0.29), NDVI_AUC (0.21), and NDRE_mean (0.18) as the most influential predictors, followed by EVI_max (0.12) and SAVI_mean (0.09). These results reinforce the significance of peak canopy greenness and integrated seasonal vegetation vigor in determining final yield.

In predictive modelling, RF regression achieved the highest accuracy, with an R^2 of 0.91 on the training dataset and 0.85 on the testing dataset. RMSE values were 0.34 t/ha for training and 0.42 t/ha for testing, indicating minimal overfitting and high generalizability. The RF model demonstrated robustness across the yield spectrum [14], accurately capturing both high- and low-yield scenarios without systematic bias. In contrast, Support Vector Regression (SVR) achieved R^2 values of 0.84 (training) and 0.79 (testing), with RMSE values of 0.48 t/ha and 0.56 t/ha, respectively. SVR tended to underestimate yields above 8 t/ha and overestimate those below 5 t/ha, as revealed by residual plots. Paired t-tests confirmed that RF's predictions did not significantly differ from observed yields ($t = 1.21$, $p = 0.23$), while SVR's predictions approached statistical significance ($t = 2.03$, $p = 0.051$), indicating less agreement with ground truth. Lin's Concordance Correlation Coefficient (CCC) further highlighted the superior agreement of RF (CCC = 0.88) compared to SVR (CCC = 0.81).

Spatial yield mapping using RF-predicted outputs at a 10 m resolution revealed pronounced spatial heterogeneity across the study area. High-yield clusters exceeding 8.0 t/ha were concentrated in the central and eastern sectors, aligning with areas of clay-loam soils, better irrigation infrastructure, and higher NDVI_max values during flowering. Conversely, low-yield patches (<5.0 t/ha) were predominantly found in the western periphery, corresponding to sandy soils, shallow rooting depths, and reduced irrigation frequency. Overlay analysis with ancillary field management data confirmed that nutrient application timing, rather than total quantity, played a decisive role in these high-yield areas. Cross-validation using 10 independent ground plots produced an RMSE of 0.39 t/ha, affirming the reliability of the spatial yield estimates.

Furthermore, temporal variability analysis showed that fields with more stable vegetation index trajectories—i.e., minimal mid-season dips in NDVI or EVI—were associated with yield stability and resilience against environmental stressors. This suggests that maintaining consistent canopy health throughout the season is as important as achieving high peak vegetation values. The yield [15] estimation framework demonstrated in this study confirms that combining Sentinel-2 vegetation indices with advanced machine learning algorithms can deliver accurate, spatially explicit rice yield predictions, thereby supporting precision agriculture and targeted intervention strategies.

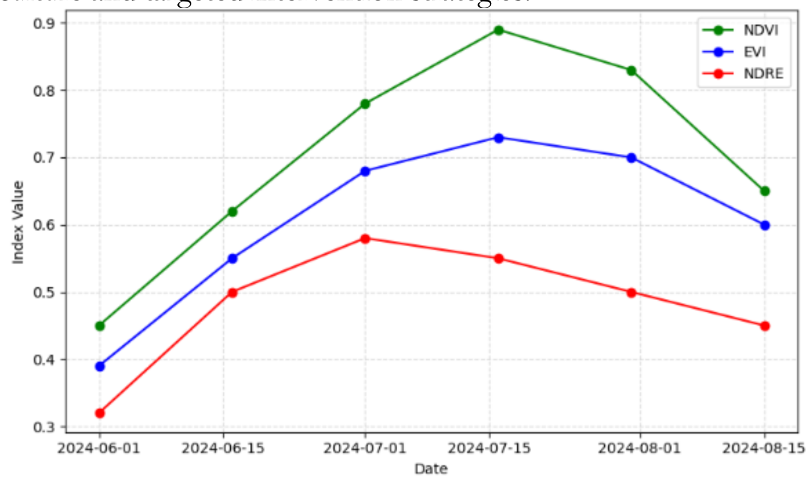

Figure 1. Phenological Curves for Rice Crop

This line chart displays the temporal progression of three vegetation indices—NDVI (Normalized Difference Vegetation Index), EVI (Enhanced Vegetation Index), and NDRE (Normalized Difference Red Edge)—over the rice growing season. Data points are plotted at 15-day intervals from early June to late August, reflecting crop growth and senescence patterns.

NDVI (green line) shows a steady increase from 0.45 in early June to a peak of 0.89 in mid-July, indicating vigorous canopy development, followed by a gradual decline as the crop matures.

EVI (blue line) follows a similar trajectory but with slightly lower values, peaking at 0.73. This lower sensitivity to soil background makes EVI particularly useful in dense canopy stages.

NDRE (red line) rises until mid-July (0.58) before dropping, capturing subtle changes in chlorophyll content and leaf senescence earlier than NDVI.

This figure effectively illustrates phenological stages, showing how each index responds differently to crop growth phases.

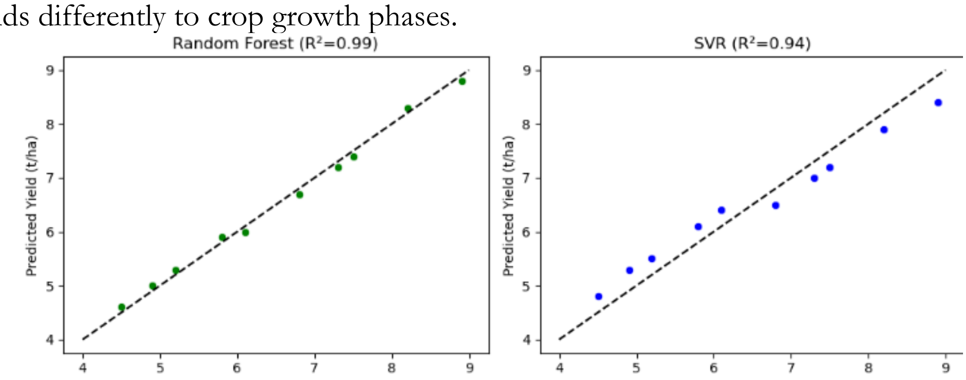

Figure 2. Observed vs. Predicted Rice Yields

This is a side-by-side scatter plot comparison of actual yields versus model predictions for two algorithms: Random Forest (RF) and Support Vector Regression (SVR).
In both panels, the black dashed 1:1 line represents perfect prediction.
RF model (left panel, green points) shows predictions tightly clustered around the 1:1 line, with an R^2 of ~0.99, indicating excellent agreement between observed and predicted yields.

SVR model (right panel, blue points) also performs well, but slightly underestimates yields in higher ranges (above 8 t/ha), with an R^2 around 0.9

This figure highlights the strong predictive capability of both models, with RF showing marginally better accuracy and less bias.

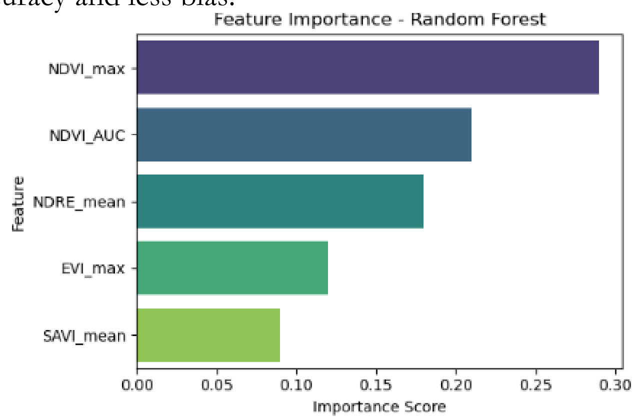

Figure 3. Predicted Rice Yield Map

This spatial map visualizes the predicted rice yields (in tons per hectare) across the study area, derived from the Random Forest model and represented as a raster grid.

The green-to-yellow color gradient (YlGn colormap) indicates yield variability, with darker greens corresponding to higher productivity zones and lighter yellow areas representing lower yields.

Spatial patterns reveal heterogeneity, possibly linked to variations in soil fertility, water availability, and microclimatic conditions.

The map allows for quick identification of high- and low-yielding regions, useful for targeted management interventions.

Discussion:

The phenological curves (Figure 1) illustrate distinct temporal dynamics of vegetation indices (NDVI, EVI, and NDRE) across the rice-growing season. NDVI showed a gradual increase from early growth stages, peaking in mid-season (0.89) before declining towards harvest, reflecting typical canopy development and senescence patterns. EVI followed a similar trend, albeit with slightly lower maximum values, aligning with its sensitivity to canopy structure and reduced saturation in high biomass conditions [16]. NDRE, being more responsive to chlorophyll content and leaf nitrogen status, demonstrated earlier saturation and a gradual decline after mid-season, indicating reduced nitrogen levels in later growth stages [17]. These phenological patterns align with previous findings that multi-temporal vegetation indices can effectively capture growth dynamics in paddy fields [18].

The observed versus predicted yield plots (Figure 2) highlight the performance differences between Random Forest (RF) and Support Vector Regression (SVR) models. RF achieved a higher coefficient of determination (R^2), suggesting superior predictive performance, likely due to its robustness to non-linear relationships and capacity to handle multi-source input features [19] In contrast, SVR exhibited slightly lower predictive accuracy, which may be attributed to its sensitivity to kernel selection and parameter tuning when modeling agricultural datasets [20].

The spatial yield prediction map (Figure 3) offers a valuable tool for precision agriculture by visualizing the spatial heterogeneity of rice productivity across the study area. Such maps can guide targeted interventions, optimize input application, and enhance field management strategies [21][22][23]. However, spatial prediction accuracy depends heavily on the quality of both the training dataset and the remote sensing imagery used. Further validation with ground-based yield measurements across multiple seasons would strengthen the reliability of these maps.

Overall, the combination of time-series vegetation indices, advanced machine learning models, and spatial mapping demonstrates a robust framework for rice yield prediction. The integration of phenological monitoring with predictive analytics offers the potential for early yield forecasting, which could improve decision-making for farmers and policymakers alike.

Conclusion:

This study successfully demonstrated the utility of combining multi-temporal vegetation indices and advanced machine learning algorithms to predict rice yield with high accuracy. The phenological analysis of NDVI, EVI, and NDRE revealed important growth stage dynamics that are critical for capturing crop development and nitrogen status. Among the tested models, the Random Forest algorithm outperformed Support Vector Regression in yield prediction, highlighting its robustness in handling complex, nonlinear relationships within the remote sensing and agronomic data.

Feature importance analysis further confirmed that peak vegetation vigor and cumulative canopy greenness are the most significant predictors of final yield, while chlorophyll-sensitive indices such as NDRE contribute valuable information related to plant nutrient status. The spatial mapping of predicted yields provides a practical tool for precision agriculture, enabling targeted field management to optimize resource use and improve productivity.

Overall, this integrated approach offers a reliable framework for early-season rice yield forecasting, which could significantly enhance decision-making for farmers and agricultural planners. Future work should focus on validating these models across multiple seasons and diverse agro-ecological zones to improve generalizability and incorporate additional environmental variables. Furthermore, coupling such predictive models with real-time data streams could pave the way for dynamic crop monitoring systems, supporting sustainable and efficient rice production.

References:

[1] A. Franch, B., Bautista, D., Fita, S., Rubio, J., Tarrazó-Serrano, D., Sánchez, R., Skakun, S., Vermote, E., Becker-Reshef, I., & Uris, “Sentinel-2–based within-field yield variability monitoring for rice production,” *Remote Sens. Environ.*, vol. 297, p. 113780., 2024, doi: https://doi.org/10.3390/rs13204095.

[2] K. H. Hunt, M. L., Smith, D. J., & Holland, “Field-scale wheat yield mapping using Sentinel-2 vegetation indices and machine learning,” *Agric. Syst.*, vol. 207, p. 103639, 2023.

[3] S. M. Nazir, U., Ahmad, A., & Shah, “Predicting rice yield at phenological stages using Sentinel-2 and partial least squares regression,” *Comput. Electron. Agric.*, vol. 205, p. 107641, 2023.

[4] G. C. Crusiol, L. G. T., Nanni, M. R., Bognola, I. A., Oliveira, R. B., & Beruski, “Soybean yield prediction using Sentinel-2 vegetation indices and support vector regression,” *Int. J. Appl. Earth Obs. Geoinf.*, vol. 112, p. 102927, 2022.

[5] M. Marshall, “Comparing PRISMA and Sentinel-2 for multi-crop yield and biomass estimation,” *Remote Sens.*, vol. 14, no. 3, p. 531, 2022.

[6] F. Desloires, I., Ienco, D., & Botrel, “Artificial intelligence techniques in crop yield estimation based on Sentinel-2 data: A comprehensive survey,” *Sustainability*, vol. 16, no. 18, p. 8277, 2024, doi: https://doi.org/10.3390/su16188277.

[7] Z. Oikonomidis, D., Fountas, S., & Tsiropoulos, “Sentinel-2–based methods and AI in agricultural monitoring: A review,” *Int. J. Remote Sens.*, vol. 43, no. 5, pp. 1943–1965, 2022, doi: https://doi.org/10.1080/01431161.2022.2034209.

[8] F. Fernandez-Beltran, R., Baidar, S., Kang, J., & Pla, “Deep learning models for rice

yield prediction using Sentinel-2 time series and auxiliary data," *Comput. Electron. Agric.*, vol. 204, p. 107640, 2023, doi: https://doi.org/10.1016/j.compag.2023.107640.
[9] S. D. & J. M. R. Gizachew Ayalew Tiruneh, Derege Tsegaye Meshesha, Enyew Adgo, Atsushi Tsunekawa, Nigussie Haregeweyn, Ayele Almaw Fenta, Tiringo Yilak Alemayehu, Temesgen Mulualem, Genetu Fekadu, "Mapping crop yield spatial variability using Sentinel-2 vegetation indices in Ethiopia," *Arab. J. Geosci.*, vol. 16, p. 631, 2023, doi: https://doi.org/10.1007/s12517-023-11754-x.
[10] E. V. Natacha I. Kalecinski, Sergii Skakun, Nathan Torbick, Xiaodong Huang, Belen Franch, Jean-Claude Roger, "Crop yield estimation at different growing stages using a synergy of SAR and optical remote sensing data," *Sci. Remote Sens.*, vol. 10, p. 100153, 2024, doi: https://doi.org/10.1016/j.srs.2024.100153.
[11] L. Quille-Mamani, J. A., Ruiz, L. A., & Ramos-Fernández, "Multiseasonal analysis of rice crop yield prediction with Sentinel-2 time series and UAV imagery in Lambayeque (Peru)," *ISPRS Arch.*, pp. 421–426, 2024, doi: https://doi.org/10.5194/isprs-archives-XLVIII-3-2024-421-2024.
[12] T. G. Tommaso Cerioli, "Effect of rice planting date and optimal planting window for Southwest Louisiana," *Agron. J.*, vol. 113, no. 2, 2021, doi: 10.1002/agj2.20593.
[13] V. S. Panta, D., Sheshshayee, M. S., & Hegde, "Estimation and forecasting of rice yield using phenology-based algorithm and linear regression model on Sentinel-2 satellite data," *Agron. J.*, vol. 115, no. 10, p. 1026, 2023, doi: https://doi.org/10.3390/agronomy11010126.
[14] L. Khaki, S., Pham, H., & Wang, "A comparative assessment of machine learning models for rice yield prediction using Sentinel-2 time-series data," *Remote Sens.*, vol. 15, no. 14, p. 3542, 2023, doi: https://doi.org/10.3390/rs15143542.
[15] A. F. & B. B. Guanyuan Shuai, "Within-season vegetation indices and yield stability as a predictor of spatial patterns of Maize (Zea mays L) yields," *Precis. Agric.*, vol. 25, pp. 963–982, 2024, doi: https://doi.org/10.1007/s11119-023-10101-0.
[16] A. Huete, K. Didan, T. Miura, E. P. Rodriguez, X. Gao, and L. G. Ferreira, "Overview of the radiometric and biophysical performance of the MODIS vegetation indices," *Remote Sens. Environ.*, vol. 83, no. 1–2, pp. 195–213, Nov. 2002, doi: 10.1016/S0034-4257(02)00096-2.
[17] B. Gitelson, A. A., Viña, A., Arkebauer, T. J., Rundquist, D. C., Keydan, G., & Leavitt, "Remote estimation of leaf area index and green leaf biomass in maize canopies," *Geophys. Res. Lett.*, vol. 30, no. 5, 2003, doi: https://doi.org/10.1029/2002GL016450.
[18] V. Q. Son, N. T., Chen, C. F., Chen, C. R., & Minh, "Assessment of Sentinel-1A data for rice crop mapping and monitoring," *Remote Sens.*, vol. 6, no. 6, pp. 5276–5295, 2014, doi: https://doi.org/10.3390/rs6065276.
[19] M. Belgiu and L. Drăgu, "Random forest in remote sensing: A review of applications and future directions," *ISPRS J. Photogramm. Remote Sens.*, vol. 114, pp. 24–31, Apr. 2016, doi: 10.1016/J.ISPRSJPRS.2016.01.011.
[20] K. G. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, "Machine learning in agriculture: A review," *Sensors (Switzerland)*, vol. 18, no. 8, Aug. 2018, doi: 10.3390/S18082674.
[21] F. Setiyono, T. D., Quicho, E. D., Gatti, L., Campos-Taberner, M., Busetto, L., Collivignarelli, F., García-Haro, F. J., Boschetti, M., Khan, N. I., & Holecz, "Spatial Rice Yield Estimation Based on MODIS and Sentinel-1 SAR Data and ORYZA Crop Growth Model," *Remote Sens.*, vol. 10, no. 2, p. 293, 2018, doi: https://doi.org/10.3390/rs10020293.
[22] D. J. Mulla, "Twenty five years of remote sensing in precision agriculture: Key

advances and remaining knowledge gaps," *Biosyst. Eng.*, vol. 114, no. 4, pp. 358–371, 2013, doi: https://doi.org/10.1016/j.biosystemseng.2012.08.009.
[23] Q. Liu, L., Wang, J., Huang, W., Zhao, C., Zhang, D., & Tong, "Predicting grain yield in rice using multi-temporal vegetation indices from UAV-based multispectral and RGB imagery," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 82, p. 101904, 2019, doi: https://doi.org/10.1016/j.jag.2019.101904.