Temporal Sentinel-2 Imagery for Wheat mapping and monitoring: Analyzing Phenological Stages with Machine Learning to Improve Mapping Precision for Small Farms

Bikesh Bade[1], Urs Schulthess[2], Gerald Blasch[3], Sonam Rinchen Sherpa[4], Him Lal Shrestha[5], Dristy Bajimaya[6], Bhavani Pinjarla[7], Mustafa Kamal[8], Sujan Nepali[9]

[1] ESRI, Sharjah, UAE, bikesbade@gmail.com
[2]CIMMYT-Turkey, Izmir, Turkey, U.Schulthess@cgiar.org
[3]CIMMYT-Mexico, Texcoco, Mexico, G.BLASCH@cgiar.org
[4]CIMMYT-Nepal, Kathmandu, Nepal, S.SHERPA@cgiar.org
[5]KAFCOL, Kathmandu, Nepal, hlshrestha@gmail.com
[6]CIMMYT-Nepal, Kathmandu, Nepal, bajimayadristy@gmail.com
[7]CIMMYT-India, Delhi, India, B.PINJARLA@cgiar.org
[8]CIMMYT-Bangladesh, Dhaka, Bangladesh, M.KAMAL@cgiar.org
[9]ICIMOD, Kathmandu, Nepal, nepsujan17@gmail.com

Keywords: Wheat Map, Small Farms, Phenological Stages, Machine Learning, Precision Mapping.

Abstract

Precise mapping and tracking of wheat crops are crucial to improve agricultural management, particularly for small farms in challenging landscapes such as Nepal. By utilizing temporal Sentinel-2 imagery, this research maps wheat fields by examining phenological stages using machine learning methods, which enhances classification accuracy. Sentinel-2, a component of the Copernicus program by the European Space Agency, offers high-quality multispectral images for precise monitoring of crop growth over time. Two classification models, Random Forest (RF) and Support Vector Machine (SVM), were employed to distinguish wheat from non-wheat areas. The accuracy of classification was improved by integrating in-situ data collected with Kobo Toolbox. The findings showed that Random Forest performed better than SVM, reaching 99% accuracy in training and 86% in validation, with 56% of the study region classified as wheat. RF's outstanding performance is due to its capacity to manage temporal and spectral intricacies, particularly in capturing the phenological cycle of crops. This research showcases how machine learning, specifically Random Forest, can enhance the accuracy of wheat mapping for small farms by analyzing phenological stages effectively, with plans to apply these methods to rice and maize in the future.

1. Introduction

Wheat is a crucial grain crop that is fundamental for global food security, offering a substantial amount of energy and protein to a sizable part of the world's population. Nevertheless, increasing worldwide food consumption because of population expansion and economic growth has resulted in challenges for wheat production (Mishra et al., 2023). In Nepal, wheat is the third most important cereal crop and plays a crucial role in rural economies, particularly during the winter Rabi season. Even though wheat cultivation is significant in Nepal, it is typically done through traditional subsistence farming, leading to lower yields because of restricted access to modern inputs, pest control, and irrigation systems and modern monitoring systems. (Qamer et al., 2014).

Recently, remote sensing and machine learning has become vital for enhancing agricultural monitoring and decision-making. Sentinel-2 satellite images, known for their detailed, multispectral features, provide an affordable, widespread monitoring option that outperforms conventional, land-based techniques. The timing of crop life cycle events, known as crop phenology, is essential for evaluating crop productivity and improving management techniques. According to Sakamoto et al. (2005), by combining phenological information with remote sensing, specifically utilizing Sentinel-2, it is feasible to effectively observe the growth phases and condition of wheat. This study aims to create a specialized approach for mapping wheat in Nepal using Sentinel-2 satellite images, with a focus on small-scale farms.

2. Study Area

The study area is located in the western Terai region of Nepal, encompassing six municipalities across Kailali and Bardiya districts. Defined by geographical coordinates ranging from 28° 21' 49'' to 28° 37' 26'' N latitude and 81° 15' 44'' to 80° 53' 33'' E longitude, this region is characterized by fertile alluvial plains, essential for agriculture. The subtropical climate features distinct wet and dry seasons, influencing cropping patterns, with rice predominating in the summer and wheat in the winter. The agricultural landscape comprises a mix of irrigated and rainfed systems, primarily managed by smallholder farmers. Understanding the region's geographical and seasonal dynamics is crucial for interpreting the outcomes of this research, which leverages temporal Sentinel-2 imagery to analyze phenological stages of wheat and enhance mapping precision for smallholder farms.

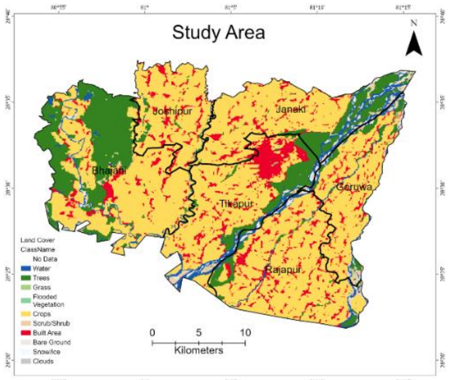

Figure 1. Western Terai region (Study Area).

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-G-2025-143-2025

3. Data and Research Methodology

The methodology for this research involves a comprehensive approach to investigate agricultural dynamics in the study area to map cropland followed by the step in the figure 2.

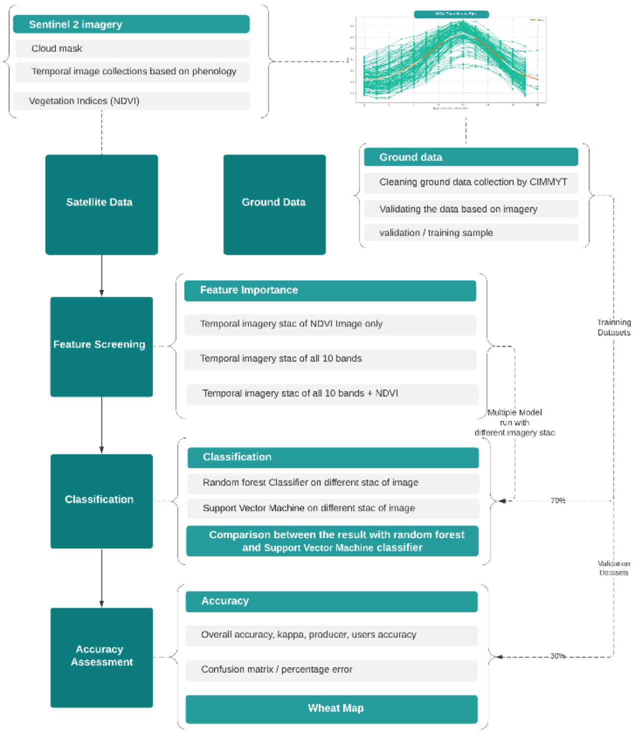

Figure 2. Methodology flow diagram.

3.1 Satellite Data

For satellite data, multi-temporal Sentinel-2 imagery is obtained, providing a comprehensive view of the study area's land cover and vegetation dynamics over time. The imagery's orbital properties for Level-2A (L2A) and Level-2B (L2B) data are carefully checked to ensure proper alignment and suitability for the analysis. Also taken in the consideration of tile and satellite orbit and its orientation. Tile T44RNS and T44RMS falls in the study area and satellite orbit is 19 and orientation is descending as represented in figure 3. Satellite Imagery with cloud cover below 10% is chosen for more reliable results. Cloud masks and

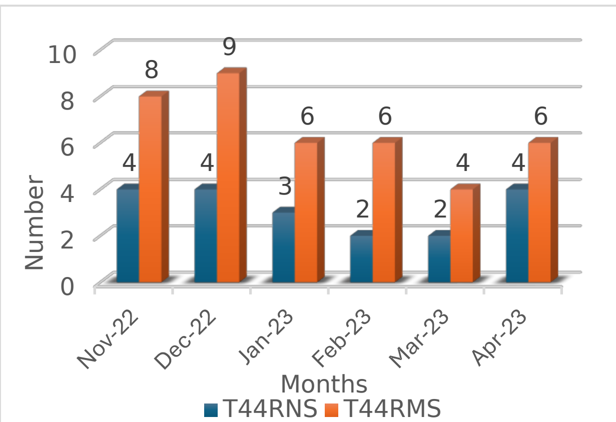

Figure 3. Satellite Data.

gap filling techniques are applied to address cloud cover and data gaps, ensuring a continuous and reliable dataset.

3.2 In- situ Data

To collect in-situ data, we conducted field surveys in collaboration with a CIMMYT Nepal Team, targeting farmlands within the study area. A structured questionnaire was administered to local farmers to gather detailed information on cropland characteristics, utilizing the Kobo Toolbox application to streamline data collection by CIMMYT field surveyors. This approach facilitated the systematic addressing of pertinent questions while accounting for the spatial layout of the crops. Following data collection, the dataset was rigorously cleaned

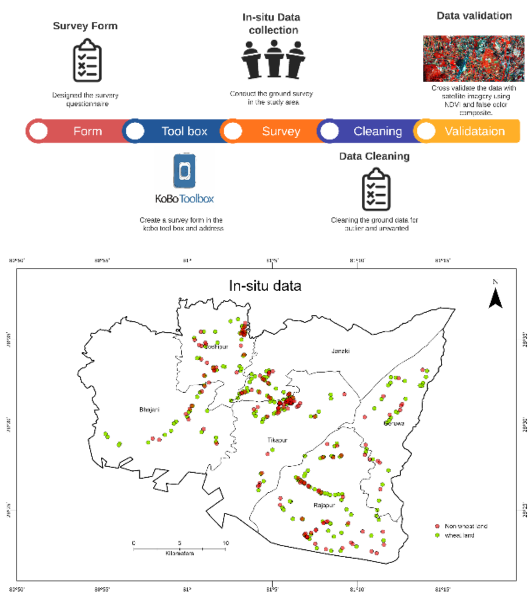

Figure 4. In-situ data map.

using ArcGIS Pro, and validation was performed by cross-referencing the survey data with satellite imagery. The methodology for ground data collection is illustrated in Figure 4.

3.3 Pre-processing and Feature Screening

The initial step in this study involved the acquisition of Sentinel satellite images spanning from November 2022 to April 2023, ensuring a comprehensive temporal coverage. Cloud masking techniques were applied, capitalizing on the advantageous meteorological conditions during the winter season, which yielded cloud-free imagery. The pre-processed imagery was then used to compute vegetation indices, such as the Normalized Difference Vegetation Index (NDVI). Crop health and condition are measured by NDVI. This vegetation index, which measures greenness, has a significant relationship with green biomass, a measure of growth. Ground truth data were then fused with satellite imagery to plot NDVI charts for all the acquired images, facilitating the construction of a phenology timeline. Plotted the NDVI chart from November 2022 to April 2023 for all the different crop in study area to get the temporal information.

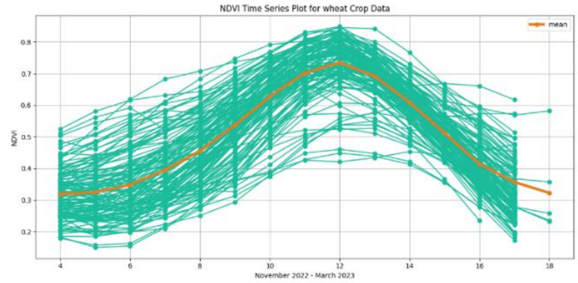

Figure 5. NDVI time series plot for wheat crop.

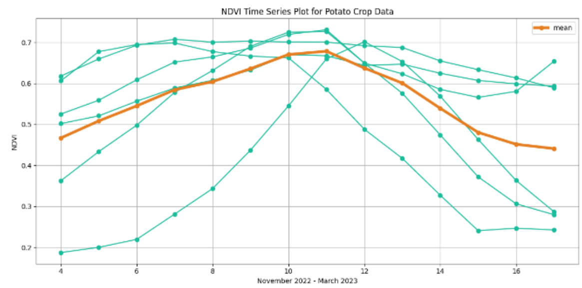

Figure 6. NDVI time series plot for Potato.

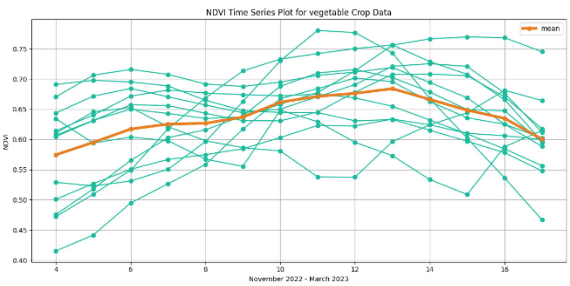

Figure 7. NDVI time series plot for vegetable.

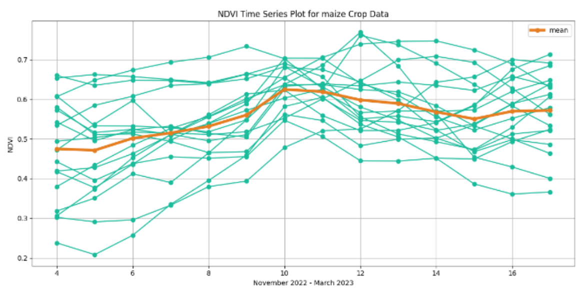

Figure 8. NDVI time series plot for Maize crop.

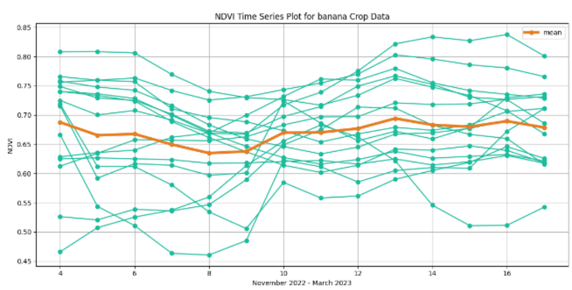

Figure 9. NDVI time series plot for Banana.

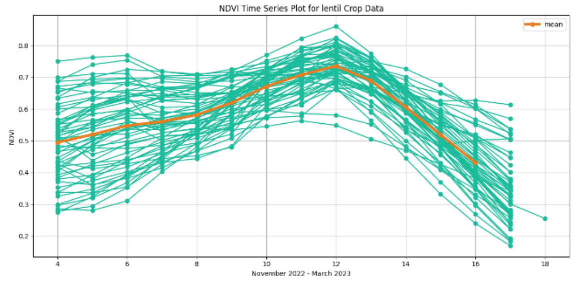

Figure 10. NDVI time series plot for lentil crop.

Figure 11. NDVI time series plot for Shrub tree.

The NDVI charts offer valuable insights into the phenological development of different crops from November 2022 to April 2023. Wheat shows a significant increase in NDVI values, starting at 0.3 in November and reaching 0.8 by February, reflecting its rapid growth during the winter season. In contrast, maize, a summer crop, experiences a slight NDVI increase from 0.4 to 0.5 as it begins growing in warmer months. Potatoes exhibit a gradual NDVI increase from 0.4 to 0.6 due to their tuberous nature and lower reliance on above-ground foliage. Lentils follow a similar pattern to potatoes, with a steady but lower NDVI increase. Shrubs demonstrate fluctuating NDVI values, decreasing slightly in winter and recovering by April, while vegetables exhibit variation based on type, with leafy vegetables such as spinach showing rapid growth compared to root vegetables. As a perennial crop, bananas maintain consistently high NDVI values throughout the period, reflecting their large leaf area index. These NDVI-derived patterns are particularly important for crop classification, especially for wheat, as its rapid and distinguishable NDVI rise during winter enables its accurate identification in satellite imagery. This data is crucial for enhancing wheat mapping accuracy, supporting precision agriculture, and improving crop dynamics monitoring using remote sensing technologies

3.4 Classification

Prior to classification, in-depth statistical analyses are carried out to pinpoint the most significant features, indices, and spectral bands found in Sentinel-2 imagery captured between November 2022 and March 2023. This procedure entails numerous rounds of model testing, with a focus on spectral bands like RED, GREEN, BLUE, NIR, and SWIR, as well as indices such as NDVI and NDWI, as depicted in the figure 12. Through the implementation of a feature screening process, the analysis ensures the consideration of temporal variations, such as crop growth cycles. This rigorous feature selection enhances the model's capacity to accurately categorize crops like wheat and maize, thereby notably enhancing the dependability of remote sensing data in agricultural monitoring.

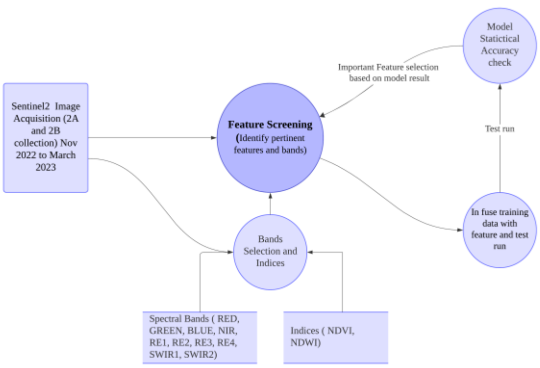

Figure 12. Feature screening chart.

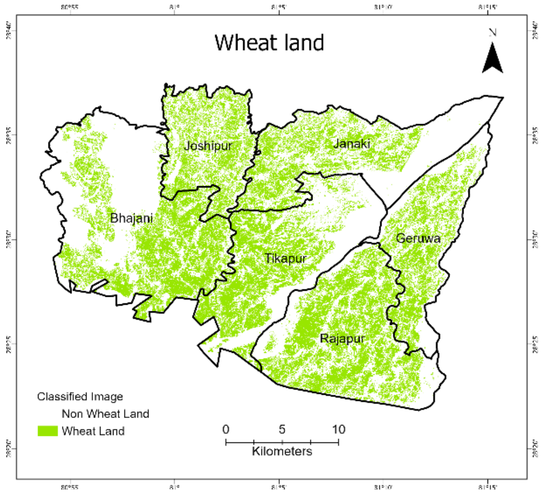

Figure 13. Random Forest resulted wheat map.

3.4.1 Random Forest Algorithm

As Table 1 illustrates, the Random Forest algorithm performed remarkably well in differentiating between wheat and non-wheat land, obtaining an overall accuracy of 99 percent. This is the breakdown of the training and validation metrics:

Table 1 RF Accuracy Table

Metric	Value
Training Accuracy	99%
Validation Accuracy	86%
Consumer Accuracy (Wheat)	87.8%
Consumer Accuracy (Non-Wheat)	86%

Scenario Analysis:

1. Four Bands (All Temporal Images): The classifier trained with red, blue, green, NIR, and NDVI resulted in a validation accuracy of 76%.

2. Ten Bands (All Temporal Images): Utilizing all 10 Sentinel-2 bands yielded a validation accuracy of 83%, with consumer accuracies of 63.6% (wheat) and 90.9% (non-wheat).

3. Four Bands (Top Ten Important Images): The model improved with an accuracy of 86%, better identifying wheat pixels (consumer accuracy of 87.8%)

4. Ten Bands (Top Ten Important Images): This scenario achieved an overall accuracy of 80.37%, with wheat consumer accuracy at 84.78%.

Random Forest classification estimated the total wheat land area to be 253.872 km², accounting for 56% of the total land area in study area.

3.4.2 Support Vector Machine (SVM)

Support Vector Machine (SVM): This algorithm achieved an overall accuracy of 90% in classifying wheat and non-wheat land effectively (Table 2). The comprehensive metrics consist of the following:

Table 2: SVM Accuracy Table

Metric	Value
Training Accuracy	90%
Validation Accuracy	72.83%
Consumer Accuracy (Wheat)	70.97%
Consumer Accuracy (Non-Wheat)	73.77%

Scenario Analysis:

1. Four Bands (All Temporal Images): The classifier trained with red, blue, green, NIR, and NDVI resulted in a validation accuracy of 72.83%.

2. Ten Bands (All Temporal Images): Utilizing all 10 Sentinel-2 bands yielded a validation accuracy of 78.95% with a consumer accuracy of 65.22% (wheat) and 84.62% (non-wheat).

3. Four Bands (Top Ten Important Images): This scenario achieved with an accuracy of 74.8% with a consumer accuracy of 67.2% (wheat) and 81.6% (non-wheat).

4. Ten Bands (Top Ten Important Images): This scenario achieved an overall accuracy of 77.8%, with wheat consumer accuracy at 78.95%.

Based on the SVM classification, the total area of wheatland is 298.02 square kilometers, while non-wheat land covers 157.69 square kilometers. This indicates that wheatland constitutes approximately 65.39% of the total land area, with non-wheat land making up about 34.60%

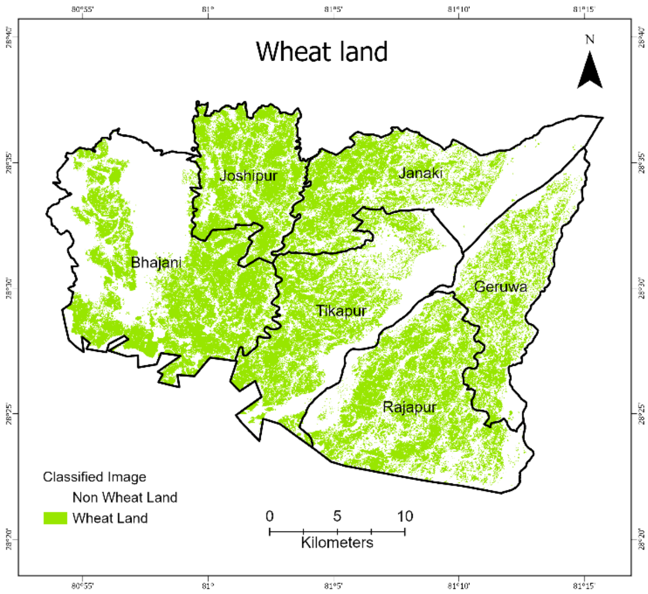

Figure 14. SVM resulted wheat map.

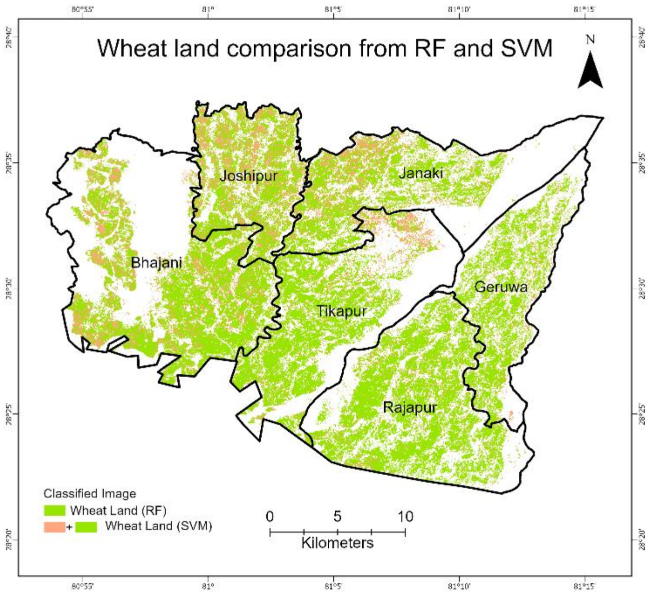

Figure 15. RF and SVM wheat land comparison map.

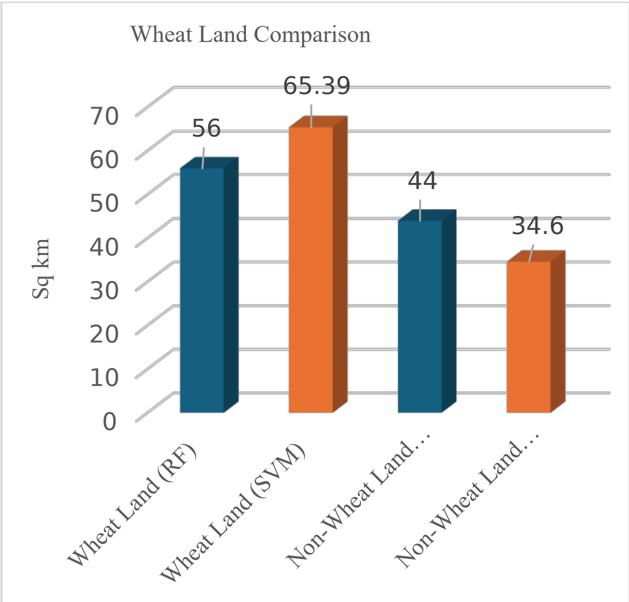

Figure 16. RF and SVM wheat land comparison chart.

4. Result and Discussion

The Random Forest (RF) algorithm showed a notable edge in classification accuracy, with an overall 86% accuracy, outperforming the Support Vector Machine (SVM) which achieved 78% accuracy in ideal conditions. Researchers can easily distinguish key features that differentiate wheat land from other land cover types using the interpretability of the RF model, which is crucial for efficient agricultural management. During scenario analyses, the model that was trained with only four bands (red, blue, green, NIR, and NDVI) achieved a validation accuracy of 76%. However, when all ten 10 Sentinel-2 bands were utilized, the accuracy increased to 83%, with wheat consumer accuracy at 63.6% and non-wheat at 90.9%. Training with the top ten crucial images using four bands notably enhanced the model's accuracy to 86%, while wheat consumer accuracy increased to 87.8%. On the other hand, the total accuracy of the top ten important images ten bands was 80.37%, with an accuracy for wheat consumers of 84.78%.

By including scenario analysis, the accuracy of these classification results is improved by simulating different situations that could influence the model's effectiveness. Researchers can improve land management strategies by examining how factors like temporal imagery and the significance of certain bands interact with each other. Although SVM is effective for high-dimensional data, its vulnerability to overfitting could hinder its performance in complicated scenarios. Therefore, it is anticipated that RF will continue to be the top choice for classifying wheat land, as it outperforms other algorithms and can easily adjust to changing agricultural conditions. This pattern is expected to persist with the expanding availability of satellite imagery data, which enables RF to provide more precise and dependable classifications.

5. Conclusion

This research assessed how well Support Vector Machine (SVM), and Random Forest (RF) algorithms classify wheat land in crop land. The findings showed that the RF algorithm performed better than the SVM algorithm, projecting that 56% (253 sq km) of the overall area consists of wheat land, with 44% identified as non-wheat land. This is consistent with the 2021 report from the Ministry of Agriculture & Livestock Development, which approximated the total wheat area to be around 240 square kilometers. The increased precision of RF is due to its resistance to noise and its ability to understand complex connections between the spectral features of various types of crops. Furthermore, the scenario analysis pointed out that classification accuracy is significantly influenced by varying band combinations and temporal images. An example is the scenario that performed the best, using the top ten crucial images, and reached an accuracy of 86%. This highlights the importance of selecting image date and bands, to improve machine learning results.

Due to the proven effectiveness of the RF algorithm in crop and land cover classification, it is advisable for agricultural planners, environmental monitors, and urban developers to utilize RF-produced crop type maps in their projects. Agricultural planners can use these maps to find the best areas for growing wheat, increasing crop yields and productivity by using specific strategies that consider soil, water availability, and climate conditions. Environmental monitors can use RF-created maps to monitor changes in land cover and pinpoint areas vulnerable to desertification or deforestation. Urban developers should utilize these maps to guarantee that new projects reduce environmental impact by steering clear of important agricultural areas and incorporating green spaces into their plans. Furthermore, the RF algorithm has the possibility to create novel applications like predicting crop yields, detecting pests and diseases, and monitoring the impacts of climate change. Stakeholders can improve by refining scenario analysis techniques and utilizing advanced analytical tools**.**

References

Ashourloo, D., Nematollahi, H., Huete, A., Aghighi, H., Azadbakht, M., Shahrabi, H. S., & Goodarzdashti, S. 2022. A new phenology-based method for mapping wheat and barley using time-series of Sentinel-2 images. *Remote Sensing of Environment*, 280. https://doi.org/10.1016/j.rse.2022.113206

Bhatta, R. D., Amgain, L. P., Subedi, R., & Kandel, B. P. 2020. Assessment of productivity and profitabilty of wheat using Nutrient Expert®-Wheat model in Jhapa district of Nepal. *Heliyon*, 6(6). https://doi.org/10.1016/j.heliyon.2020.e04144

Breiman, L. Random Forests. Machine Learning 45, 5–32 2001. https://doi.org/10.1023/A:1010933404324

Evgeniou, T., & Pontil, M. 2001. Support vector machines: Theory and applications. *Lecture Notes in Computer Science* (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2049 LNAI, 249–257. https://doi.org/10.1007/3-540-44673-7_12

Feng, S., Zhao, J., Liu, T., Zhang, H., Zhang, Z., & Guo, X. 2019. Crop Type Identification and Mapping Using Machine Learning Algorithms and Sentinel-2 Time Series Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(9), 3295–3306. https://doi.org/10.1109/JSTARS.2019.2922469

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202, 18–27. https://doi.org/10.1016/j.rse.2017.06.031

Guo, Y., Xia, H., Pan, L., Zhao, X., Li, R., Bian, X., Wang, R., & Yu, C. 2021. Development of a new phenology algorithm for fine mapping of cropping intensity in complex planting areas using sentinel-2 and google earth engine. *ISPRS International Journal of Geo-Information*, 10(9). https://doi.org/10.3390/ijgi10090587

Hunt, M. L., Blackburn, G. A., Carrasco, L., Redhead, J. W., & Rowland, C. S. 2019. High resolution wheat yield mapping using Sentinel-2. *Remote Sensing of Environment*, 233. https://doi.org/10.1016/j.rse.2019.111410

Joshi, B., Mudwari, A., & Bhatta, M. 1970. Wheat Genetic Resources in Nepal. *Nepal Agriculture Research Journal*, 7, 1–10. https://doi.org/10.3126/narj.v7i0.1859

Khadka, K., Torkamaneh, D., Kaviani, M., Belzile, F., Raizada, M. N., & Navabi, A. 2020. Population structure of Nepali spring wheat (Triticum aestivum L.) germplasm. *BMC Plant Biology*, 20(1). https://doi.org/10.1186/s12870-020-02722-8

Li, G., Cui, J., Han, W., Zhang, H., Huang, S., Chen, H., & Ao, J. 2022. Crop type mapping using time-series Sentinel-2 imagery and U-Net in early growth periods in the Hetao irrigation district in China. *Computers and Electronics in Agriculture*, 203. https://doi.org/10.1016/j.compag.2022.107478

Li, Y., Qin, Y., Ma, L., & Pan, Z. 2020. Climate change: vegetation and phenological phase dynamics. *International Journal of Climate Change Strategies and Management*, 12(4), 495–509. https://doi.org/10.1108/IJCCSM-06-2019-0037

Mishra, B., Bhandari, R., Bhandari, K. P., Bhandari, D. M., Luintel, N., Dahal, A., & Poudel, S. 2023. High-Resolution Mapping of Seasonal Crop Pattern Using Sentinel Imagery in Mountainous Region of Nepal: A Semi-Automatic Approach. *Geomatics*, 3(2), 312–327. https://doi.org/10.3390/geomatics3020017

Muruganantham, P., Wibowo, S., Grandhi, S., Samrat, N. H., & Islam, N. 2022. A Systematic Literature Review on Crop Yield Prediction with Deep Learning and Remote Sensing. *Remote Sensing* (Vol. 14, Issue 9). MDPI. https://doi.org/10.3390/rs14091990

Pan, Z., Huang, J., Zhou, Q., Wang, L., Cheng, Y., Zhang, H., Blackburn, G. A., Yan, J., & Liu, J. 2015. Mapping crop phenology using NDVI time-series derived from HJ-1 A/B data. *International Journal of Applied Earth Observation and Geoinformation*, 34(1), 188–197. https://doi.org/10.1016/j.jag.2014.08.011

Phiri, D., Simwanda, M., Salekin, S., Nyirenda, V. R., Murayama, Y., & Ranagalage, M. 2020. Sentinel-2 data for land cover/use mapping: A review. *Remote Sensing* (Vol. 12, Issue 14). MDPI AG. https://doi.org/10.3390/rs12142291

Piragnolo, M., ... Grigolato, S., 2021. Responding to Large-Scale Forest Damage in an Alpine Environment with Remote Sensing, Machine Learning, and Web-GIS. *Remote Sensing* 13, 1541. https://doi.org/10.3390/rs13081541

Qamer, F. M., Shah, S. N. P., Murthy, M. S. R., Baidar, T., Dhonju, K., & Hari, B. G. 2014. Operationalizing crop monitoring system for informed decision making related to food security in Nepal. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 40(8), 1325–1330. https://doi.org/10.5194/isprsarchives-XL-8-1325-2014

Sakamoto, T., Yokozawa, M., Toritani, H., Shibayama, M., Ishitsuka, N., & Ohno, H. 2005. A crop phenology detection method using time-series MODIS data. *Remote Sensing of Environment*, 96(3–4), 366–374. https://doi.org/10.1016/j.rse.2005.03.008

TAO, J., ZHANG, X., WU, Q., & WANG, Y. 2022. Mapping winter rapeseed in South China using Sentinel-2 data based on a novel separability index. *Journal of Integrative Agriculture*. https://doi.org/10.1016/j.jia.2022.10.008

Xie, Y., Sha, Z., & Yu, M. 2008. Remote sensing imagery in vegetation mapping: a review. *Journal of Plant Ecology*, 1(1), 9–23. https://doi.org/10.1093/jpe/rtm005ssss

Zhao, Y., Wang, X., Guo, Y., Hou, X., & Dong, L. 2022. Winter Wheat Phenology Variation and Its Response to Climate Change in Shandong Province, China. *Remote Sensing*, 14(18). https://doi.org/10.3390/rs14184482

Acknowledgements

We sincerely thank the CIMMYT Nepal team and KAFCOL for their support and assistance in this research. We also appreciate the valuable guidance of senior scientists, whose expertise has been instrumental in shaping this study. Additionally, we acknowledge the funding and resources provided by CIMMYT Nepal, which made this work possible.