Sentinel-2-Based Crop Classification in Rayong Province, Thailand: A Regional Remote Sensing Approach

Punyanuch Kajornchaikitti[1], Karun Tancharoen[1], Noppharoot Bavornkijsutee[1]
Totsanat Rattanakaew[2] Boonyarit Changaival[1,*]
[1]Department of Computer Engineering, King Mongkut's University of Technology Thonburi, Thailand
[2]Land Use Policy and Planning Division, Department of Land Development, Ministry of Agriculture and Cooperatives, Thailand

***Abstract*— Crop classification is essential for effective agricultural monitoring and land-use management. In Thailand, the availability of up-to-date land-use maps remains limited due to budget constraints and the traditional field surveys, which are laborious and time-consuming. This study presents an applied regional case study utilizing multispectral Sentinel-2 imagery and a Random Forest classifier to operationalize cost- and time-effective crop classification. Focused on Rayong Province, Thailand, this implementation supports the generation of localized, crop-focused land-use maps. Spectral indices and bands, including the Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), Normalized Difference Water Index (NDWI), MERIS Terrestrial Chlorophyll Index (MTCI), and the Shortwave infrared (SWIR) band, were employed to characterize crop vegetation greenness and water content, and water bodies. The proposed method was applied to classify 13 major economic crops in Rayong Province. The model attained an overall F1-score of 0.71, with oil palm achieving the highest F1-score of 0.81 among the economic crops, and the reservoir accomplished the highest F1-score of 0.94 among all classes.**

Keywords- Crop Classification, Remote Sensing, Multispectral, Random Forest, Machine learning

I. INTRODUCTION

Remote sensing has emerged as a transformative tool for the large-scale monitoring of agricultural landscapes, offering a high-frequency alternative to traditional surveying. In tropical regions like Thailand, the European Space Agency's (ESA) Sentinel-2 mission is particularly valuable, as its available multispectral data provides the high spatial and temporal resolution necessary to distinguish complex crop patterns. These technologies facilitate accurate crop mapping, a fundamental requirement for reliable yield estimation, strategic agricultural planning, and the advancement of sustainable land management. Nevertheless, the urgency for such precision is underscored by recent economic shifts. Data from the Ministry of Agriculture reveals that Thailand's agricultural price index rose steadily from 2018 to 2024, signaling a robust and growing demand for crop products [1]. While the total area of agricultural holdings expanded significantly between 2013 and 2023, the agricultural production index remained largely stagnant during the same period [1]. This discrepancy, where land expansion fails to yield a proportional increase in output, suggests a critical gap in land-use efficiency. Without detailed spatial data to identify which crops are being grown and where resources are being underutilized, Thailand struggles to optimize its existing agricultural footprint. Despite this clear need, official mapping efforts are hindered by severe institutional bottlenecks. The Land Development Department (LDD), tasked with maintaining these records, operates under significant fiscal constraints; currently, only 1.06% of their allocated budget is dedicated to map development and updates [2]. This funding gap slows the traditional field surveys, which require extensive personnel, transportation, and living expenses, making it unsustainable. As a result, national land-use data often becomes outdated before it can be effectively implemented. To address these institutional challenges, this study serves as an applied implementation of an automated, cost-effective framework for regional crop classification. By applying a Random Forest classifier to Sentinel-2 imagery, the official LDD geospatial data is integrated with a suite of spectral indices focused on vegetation greenness and water content, together with a spectral band, including the Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), Normalized Difference Water Index (NDWI), the MERIS Terrestrial Chlorophyll Index (MTCI), and the Short Wave Infrared (SWIR). Using Rayong Province, a key economic zone in eastern Thailand, as our primary case study is to classify 13 major crops such as rice, cassava, pineapple, para rubber, oil palm, durian, rambutan, coconut, mango, longan, jackfruit, mangosteen, and langsat. [3]–[5]. Through this regional implementation, this study aims to provide the LDD with a validated, scalable model to produce high-resolution, crop-focused maps that support both economic growth and environmental preservation, and enable the LDD to send staff only to conduct fieldwork on areas that require further inspection rather than all provinces to reduce the extended costs such as travel and accommodation costs as well as the time associated with traditional surveys in the long term.

This paper covers six sections: Section two reviews the literature related to this study; Section three describes the methodology, including the details of the study area, data sources, data preprocessing, dataset, spectral indices and bands, and Random Forest classification; Section four reports the results; Section five discusses the findings; and lastly, Section six provides the conclusion and outlines future work.

II. LITERATURE REVIEW

Remote sensing has become a practical alternative to traditional field surveys for creating large-scale land-use maps and crop-mapping. Its implementation with multispectral satellite imagery enables the repeated observation of crop characteristics in extensive areas at relatively low cost through the multispectral bands and extends to the derivation of vegetation indices. Sentinel-2 is one of the them which can be used for conditions that require detailed information and temporal monitoring [6], such as under traditional cultivation practices, where agricultural areas are fragmented and multiple crop species are often cultivated within the same field [6]–[8], as it contains high-resolution reflectance data at 10 m with a high revisit rate, approximately 5 days [6]. It has 13 spectral bands in total, consisting of four bands at 10 m resolution (Blue, Green, Red, and Near-Infrared (NIR)), six bands at 20 m resolution (Red Edge 1, Red Edge 2, Red Edge 3, Narrow NIR, SWIR 1, and SWIR 2), and three bands at 60 m resolution (Coastal aerosol, Water vapour, and Cirrus) [9]. These bands can be derived into various spectral indices to quantify crop characteristics and classify them, such as vegetation greenness and water content.

The widely used vegetation index of greenness is NDVI [10], which can be implemented to represent the presence and greenness of crops [11]. Previous studies suggest implementing EVI alongside NDVI, as EVI excels in dense areas, reducing the saturation effect commonly observed in NDVI [10] [11]. NDVI itself also has a problem with the atmospheric effect of clouds, where the MTCI is found to be less sensitive than NDVI due to the computation between two nearby bands, using Red Edges [12] [13]. The importance of an NDVI, EVI, and MTCI combination, surpassing dense canopy conditions through EVI and atmospheric effects through MTCI, and having NDVI as the baseline for sparse areas, shall be investigated further in this paper.

Beyond vegetation greenness indices, crop water content is also considered as a factor to classify crops, both for crop species and land use. The previous study has shown through spectroscopy and gravimetric validation experiments that the SWIR band's water absorption features exhibit a strong linear relationship with leaf water content and photosynthetic capacity across diverse crop species [14]. It has also been reported that SWIR contributes substantially to crop classification performance, accounting for approximately 0.35 of the overall feature importance when classifying major crops such as rice, corn, soybean, and other crop types [15]. To separate land and water bodies, NDWI, which uses the green and NIR bands, has shown strong performance [16]. It takes advantage of the high reflectance in the green band and the minimal reflectance in the NIR band for water, which is opposite to terrestrial vegetation and bare soil [16]. Following the standard formulation, NDWI is derived from GREEN and NIR reflectance using a normalized difference approach [16]. Thus, the characteristics can be clearly classified: when the target is water (meaning GREEN is higher than NIR), the value becomes positive, while it becomes negative for terrestrial vegetation and close to zero for bare soil [16]. These findings support the integration of SWIR, with its ability in internal water detection of crops, and NDWI, with its ability to separate water bodies from land, alongside vegetation indices, suggesting enhanced crop separability in heterogeneous agricultural landscapes.

To effectively utilize these spectral bands and derived indices for crop classification, recent studies have increasingly adopted artificial intelligence approaches, including traditional machine learning (ML) and deep learning (DL) methods. Traditional ML methods are recognized for their robustness when encountering limited or imbalanced datasets and operating on relatively low computational resources; however, they typically rely on manual feature engineering to achieve optimal performance [17]. In contrast, DL methods are suitable for large-scale datasets and can automatically learn optimized feature representations from data without the need for manual feature engineering, but they require higher computational resources [17]. Previous studies have demonstrated that traditional ML algorithms, including Random Forest (RF), Support Vector Machines (SVM), and XGBoost, exhibit strong robustness in terms of computational efficiency, such as execution time, while achieving classification performance comparable to DL models, such as Temporal Convolutional Neural Networks (TCNN), when evaluated under identical CPU and memory constraints [18].

From an operational perspective, considerations beyond classification accuracy, such as computational efficiency, interpretability, and long-term maintainability, are significant for governmental applications, notably with budget constraints and limited training samples for certain crop classes. Hence, traditional ML methods remain widely used in remote sensing classification because of their robustness and relatively low computational cost [17]. Among machine learning classifier methods, the Random Forest algorithm has been reported as effective for crop classification due to its strong generalization ability, robustness to overfitting compared to other machine learning classifier methods [4] [19]. Thus, to support the LDD of Thailand, this study aims to implement a Random Forest classifier to develop a time- and cost-effective solution for classifying 13 major economic crops in Thailand, scoped in Rayong, as shown in the Fig. 1.

III. Methodology

A. Study Area

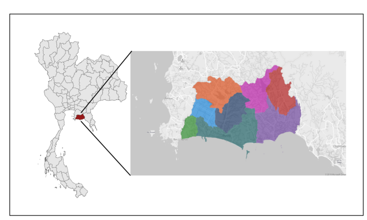

Fig. 1. Province map of Rayong, Thailand (adapted from [20] [21]).

This study was undertaken in Rayong Province, which lies approximately at 12°40'32"N and 101°16'42"E, covering a total area of 3,666 km^2 in eastern Thailand. It is selected as all 13 economic crops in Thailand are present to study, including rice, cassava, pineapple, para rubber, oil palm, durian, rambutan, coconut, mango, longan, jackfruit, mangosteen, and langsat [5].

B. Data Source

The data used in this study comprises satellite imagery and provincial land-use shapefiles in Rayong Province. These sources allow for a comparative analysis across two distinct timeframes (2020 and 2024), focusing on the blooming seasons critical for crop identification.

1) Sentinel-2 Level-2A imagery obtained from the Copernicus platform under the European Commission and the European Space Agency [6]. The imagery from October, November, and December 2020 and 2024 was used as these months correspond to the period when fruit crops bloom, making them suitable for classifying crop types and for the model to learn. To ensure the most complete dataset, a monthly composite image was generated for each month to account for cloud cover, haze, or fog.
2) A 2020 and 2024 land-use shapefile of Rayong Province, provided by the LDD of Thailand, was used in this study. The dataset contains detailed polygon features representing land-use categories across the province, including forest areas, agricultural lands labeled by crop type, and water bodies.

In summary, this study integrates Sentinel-2 Level-2A monthly composite imagery and LDD provincial land-use shapefiles from 2020 and 2024 during October to December, to support crop classification in Rayong Province.

C. Preprocessing and Data Mapping

The Sentinel-2 satellite images were preprocessed to create clean and consistent multispectral data for further analysis. The widest period available within each month was used. Relevant spectral bands and metadata were then extracted, which included B02, B03, B04, B05, B06, B07, B08, B8A, B11, B12, and the scene classification layer (SCL) refers to the standard Sentinel-2. Radiometric preprocessing was performed by applying the bottom-of-atmosphere (BOA) offset derived from the metadata, while pixel values were retained in scaled integer form. Pixels affected by clouds or invalid conditions were removed using the Sentinel-2 validity mask (VAL). Images collected within the same period were composited by calculating the median value for each pixel and band to reduce noise and remaining cloud effects. The resulting bands were stacked and saved as a multispectral GeoTIFF.

Comparison with the land-use shapefile and Sentinel-2 imagery reveals that the land-use polygons extend beyond the actual crop boundaries, potentially introducing noise into the vegetation index and consequently the classification model. To mitigate this issue, each polygon boundary was contracted inward by 30 m before rasterization [3]. Industrial areas were excluded, as land use in these regions remains largely unchanged over time [3]. The rasterized labels were then overlaid with Sentinel-2A imagery, and only the intersecting pixels were extracted for analysis. These pixels were then used to compute the vegetation indices and were spatially matched with their label for model training.

D. Dataset

The training data were primarily derived from the 2024 dataset from October to December, from which all thirteen classes corresponding to major economic crop types were extracted through preprocessing of the 2024 land-use data and its associated shapefile, as described previously. The economic crops include rice, cassava, pineapple, para rubber, oil palm, durian, mango, jackfruit, coconut, mangosteen, longan, rambutan, and langsat. However, because some crops had limited sample availability in 2024, additional samples for smaller classes (rambutan, coconut, mango, longan, jackfruit, mangosteen, and langsat) were supplemented using the 2020 dataset. These supplementary samples were processed using the same preprocessing workflow as the 2024 data to ensure consistency. In addition to the crop classes, two extra classes, reservoirs and others, were included to allow the model to distinguish water bodies and non-target land covers from the crop types of interest.

The dataset consists of 3,039,469 samples from all classes, as detailed in Table I, and was split into an 80:10:10 ratio for training, validation, and testing. This resulted in 2,431,575 samples for training, 303,947 for validation, and 303,947 for testing.

E. Spectral Indices and Bands

The formulas for all spectral indices can be viewed in Table II. This study employs 4 spectral indices and 1 spectral band to be used as the features for training the model. The spectral indices include NDVI, EVI, NDWI, and MTCI. The spectral band SWIR. These spectral features focus on vegetation greenness and water content.

- NDVI is used to represent vegetation greenness [10]. In particular, most plants strongly absorb red light, with a wavelength of about 0.63 to 0.69 micrometers, through their chlorophyll for photosynthesis purposes. In contrast, near-infrared, or NIR, with a wavelength of around 0.76 to 0.9 micrometers, is strongly reflected.

TABLE I
NUMBER OF PIXELS PER CLASS

Class	Number of pixels
Rice	156,561
Cassava	387,572
Pineapple	386,559
Para rubber	385,546
Oil palm	389,900
Durian	393,116
Mango	56,149
Jackfruit	45,192
Coconut	25,077
Mangosteen	26,568
Longan	27,900
Rambutan	7,813
Langsat	1,956
Reservoir	374,249
Others	375,311

TABLE II

VEGETATION INDICES DERIVED FROM SENTINEL-2 IMAGERY

Vegetation Index	Formula	Ref.
Normalized Difference Vegetation Index (NDVI)	$\frac{B8-B4}{B8+B4}$	[10]
Enhanced Vegetation Index (EVI)	$2.5 \cdot \frac{B8-B4}{B8+6 \cdot B4-7.5 \cdot B2+1}$	[11]
Normalized Difference Water Index (NDWI)	$\frac{B3-B8A}{B3+B8A}$	[16]
MERIS Terrestrial Chlorophyll Index (MTCI)	$\frac{B6-B5}{B5-B4}$	[12]

This means that healthy vegetation with high greenness usually has low red and high NIR values when measuring with sensors. On the other hand, the values are closer for unhealthy vegetation with low greenness. NDVI utilizes this efficiently by taking the difference between NIR and red. Hence, the word difference in the name. The problem that entails is that the difference is sensitive to lighting conditions, which means the surface imaged at a different time of day gives off a different value. NDVI solves this by dividing that difference by the sum of NIR and red. Hence, the word normalized in the name. This works because the difference and sum scale the same way, so the ratio stays stable. Overall, using the greenness detection property that NDVI has, we can distinguish between plants with different greenness, and also between plants and soil, water, and non-vegetated surfaces with no greenness.

- EVI is used to help address some of the problems with NDVI [11]. Particularly, as vegetation gets denser, red becomes increasingly smaller, whereas NIR becomes increasingly larger until a point where red gets so close to 0 that NDVI maxes out at 1. In other words, although the surface still changes meaningfully, NDVI stops changing accordingly. This is called saturation. Another problem is that soil reflects both red and NIR, often in a correlated way. On top of that, the atmosphere is also problematic. In particular, water vapor, dust, smoke, and other particles in the air either absorb or scatter light. These soil and atmospheric properties introduce bias into NDVI measurements. By including the additional parameters in the formula, EVI effectively helps reduce the severity of these problems. Overall, EVI is used here, since when these problems are present, it distinguishes better.
- NDWI is used to represent water [16]. Particularly, the reason for the formula and how it works is the same as NDVI, which is explained earlier, replacing only the NIR band in NDVI with the green band and the red band with the narrow NIR band. This is because water does not absorb or reflect light the same way as vegetation. Instead, water absorbs NIR strongly, while strongly reflecting green, with a wavelength of approximately 0.51 to 0.58 micrometers. Still, as mentioned earlier, narrow NIR is chosen here, even if NIR can also be used, since the absorption is even stronger. Overall, NDWI with this parameter choice improves the model's potential to separate water bodies from land surfaces.
- MTCI is used to represent chlorophyll content [12]. Specifically, the red edge, with a wavelength of 0.68 to 0.75 micrometers, lies in a region where light absorption in a plant starts to get weaker, which means reflection starts to get stronger. Therefore, as the chlorophyll content gets higher, absorption decays more slowly, and the red edge gets shifted to a longer wavelength, nearer to where NIR is. By taking the ratio of the difference between NIR and the red edge, and between the red edge and red, MTCI can estimate the chlorophyll content. By placing the formula like this, MTCI does not face the saturation problem, even when the vegetation in the pixel gets dense. Overall, by going into the chlorophyll content level, not just how green the pixel is, MTCI's ability to distinguish between plants with different greenness is better than NDVI, especially when NDVI saturates.
- SWIR is used to represent plant water [22]. In particular, water strongly absorbs SWIR, which helps distinguish between plants with different water content inside their fields or even inside themselves, and between water bodies and land surfaces. Overall, this spectral band can be very useful when economic crops, such as rice fields with water inside, are present, which is the case in this study.

In conclusion, the main feature of NDVI that is being used in this study is to distinguish between plants and non-vegetation surfaces. However, when it saturates or bias is introduced, EVI is used to improve data robustness and reliability. As for NDWI, it is used to separate water bodies from land surfaces. With the chlorophyll content detection property that helps distinguish between plants with different greenness better than NDVI without saturation, MTCI is used for that. Lastly, SWIR is used to distinguish between plants with different water content in their fields or themselves.

F. Random Forest Classification

Random forest is an ensemble of trees, where each tree is trained on a random subset of the dataset with replacement. For each tree and each of its splits, a random subset of features is considered [4]. The final prediction is typically obtained through a majority vote of all trees for classification, and an average for regression. In this study, a random forest is used to classify the 13 economic crops of the Rayong province of Thailand, as well as reservoirs, and all other land covers will be classified as a class called others [5].

The selected hyperparameters are shown in Table III. They were conducted to optimize the model performance using a random search with 5-fold cross-validation over 50 iterations.

IV. Results

The validation and test performance of the Random Forest Model are shown in Table IV. The Random Forest classifier reached an accuracy of 0.717, a Cohen's Kappa coefficient of 0.679, and an F1-score of 0.716, with precision and recall values of 0.727 and 0.717, respectively, on the validation set. The model also got similar performance on the test set.

As detailed in Table V, the model demonstrated strong performance in distinguishing classes with unique spectral signatures, such as reservoir (F1-score: 0.94) and oil palm (F1-score: 0.81). However, the confusion matrix reveals significant misclassification among crops with similar growth profiles, most notably pineapple and cassava, and between durian and para rubber. These errors likely stem from high spectral and radiometric overlap between crops sharing similar chlorophyll density and morphological characteristics.

V. Discussion

While the Random Forest classifier demonstrated strong capability in identifying major economic classes, the significant misclassification among certain fruit crops requires further examination. The extreme numerical disparity in our dataset, ranging from 393,116 samples for durian to merely 1,956 for langsat, presents a severe limitation. When this numerical scarcity is combined with the high spectral and structural overlap of mature tropical canopies, the model lacks sufficient diverse examples to establish clear decision boundaries for the minority classes, leading to confusion with geometrically similar crops like pineapple and cassava.

Even though oversampling techniques, such as SMOTE, ADASYN, and BorderlineSMOTE, have been implemented to tackle the imbalanced dataset, all metrics are different by no more than 0.1 compared to the models that were trained on the original imbalanced dataset when the cross-year evaluation is implemented. This consequence suggests that the model trained on the synthetic data is overfit to the samples of the synthetic data in the training year, rather than capturing generalizable crop characteristics.

Pineapple and cassava are two of the crops that the model is most confused about. They have a large overlap in NDVI, EVI, and MTCI distributions across October to December. The data suggests similar canopy development with resemblance in Leaf Area Index (LAI) and chlorophyll concentrations at the 10-m pixel scale. Water-related features provide greater discrimination. The peak NDWI of cassava is concentrated around -0.5 to -0.6, while pineapple shows a wider, flatter distribution and develops a bimodal pattern in December. For raw SWIR, the pineapple consistently has a peak difference of around 0.02 to 0.03 lower than cassava, and is most distinct in December, with approximately 0.1. The distinction of the distribution in December of SWIR and NDWI can stem from the crop response to Thailand's winter season. As a result, December is the most discriminative month in this window to separate between pineapple and cassava.

Based on feature importance as shown in Fig. 2, EVI in October plays a significant role in model prediction. October coincides with the rainy season, when crop canopies become dense due to rapid leaf growth. Because EVI remains sensitive under dense canopy conditions, where NDVI saturates, it better captures variations in canopy structure and vegetation vigor among crop types, see III-E. Consequently, crop separability is enhanced during this period, when most crops reach peak leaf development, providing strong classification performance in a heterogeneous agricultural landscape. SWIR from November and December also appears to play a critical role in the crop classification. This may be explained by seasonal changes during the early winter period, when crop water status becomes more distinct across crop types. Since SWIR is highly sensitive to leaf and canopy water content, it can capture differences in water retention

TABLE III
Selected Hyperparameters from Random Search with 5-Fold Cross-Validation

Hyperparameter	**Value**
Class weight	balanced
Maximum tree depth	36
Max features	$\log_2$
Minimum samples per leaf	1
Minimum samples to split	5
Number of trees ($n_{\text{estimators}}$)	254

TABLE IV
Validation and Test Performance of the Random Forest Model

Metric	**Validation**	**Test**
Accuracy	0.7174	0.7157
Kappa	0.6795	0.6775
F1-score	0.7155	0.7139
Precision	0.7273	0.7260
Recall	0.7174	0.7157

TABLE V
Test Performance on Each Economic Crop Type

Class	**Precision**	**Recall**	**F1-score**	**Support**
Rice	0.68	0.80	0.74	15,656
Cassava	0.63	0.71	0.67	38,758
Pineapple	0.69	0.67	0.68	38,656
Para rubber	0.72	0.73	0.72	38,555
Oil palm	0.83	0.79	0.81	38,990
Durian	0.57	0.74	0.65	39,311
Rambutan	0.93	0.20	0.33	781
Coconut	0.76	0.34	0.47	2,507
Mango	0.76	0.53	0.62	5,615
Longan	0.87	0.54	0.67	2,790
Jackfruit	0.77	0.45	0.57	4,519
Mangosteen	0.79	0.31	0.44	2,657
Langsat	0.76	0.33	0.46	196
Reservoir	0.95	0.93	0.94	37,425
Others	0.69	0.54	0.61	37,531

characteristics. All NDWI features from October, November, and December rank among the most important predictors, as they effectively distinguish crops, suggesting that they are particularly useful for identifying rice, which is typically cultivated in flooded fields, compared with other crops and terrestrial vegetation.

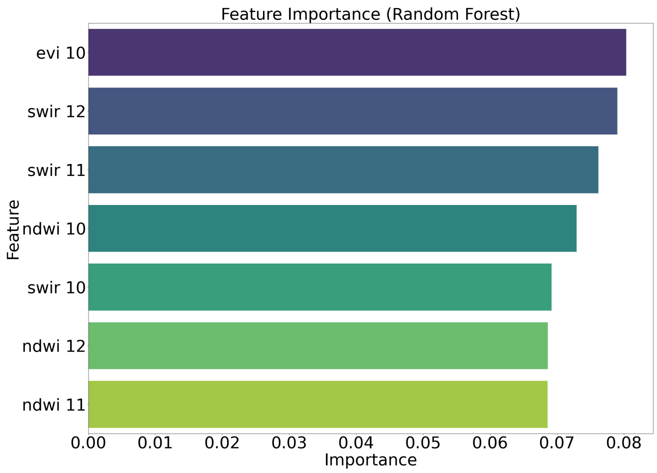

Fig. 2. Seven most contributing features across all features for each month: October (10), November (11), and December (12).

VI. Conclusion

This regional case study aims to implement a Random Forest classifier to develop a time- and cost-effective solution for classifying 13 major economic crops in Thailand. As a result, the implementation excelled at identifying distinct classes like oil palm and reservoirs, but struggled to separate other crops, including cassava and pineapple, due to spectral similarities, limited training samples, and inaccuracies in the reference data. The analysis also reveals that December provides the most discriminative information of the water-related features, such as NDWI and SWIR, for the cassava and pineapple separation. Further study can investigate more of the water-related features.

To improve this framework for operational use, future implementations should integrate Sentinel-1 Synthetic Aperture Radar (SAR) with Sentinel-2 imagery. SAR provides essential structural context to differentiate morphologically similar crops. Incorporating a Digital Elevation Model (DEM) can supply topographic predictors (like elevation and slope) for crops with specific terrain preferences. Expanding the observation period beyond the current October to December window can capture complete growth cycles and field preparation stages, establishing a robust operational baseline for heterogeneous tropical landscapes.

References

[1] National Statistical Office, “Thai economic indicators 2025,” 2025, accessed: Jan. 8, 2026. [Online]. Available: https://nso.go.th/public/e-book/Indicators-Economic/Thai_Economic_Indicators_2025/48/

[2] Land Development Department, Ministry of Agriculture and Cooperatives, “Budget report on undertaking operations 2025,” accessed: Jan. 8, 2026. [Online]. Available: https://orgweb.ldd.go.th/sources/img/part_document/FILES_12996_68426b206365e.pdf

[3] Land Development Department, Ministry of Agriculture and Cooperatives, “Present land use map (1:25,000) shapefile: Rayong province,” 2024, accessed: Jan. 9, 2026. [Online]. Available: https://lddcatalog.ldd.go.th/dataset/ldd_21_01/resource/c62118eb-c88d-4639-8593-3e60c48d09cc?view_id=4d8fab72-eb4f-47e7-bdbb-1bc4a502b8d7

[4] L. Breiman, “Random forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[5] Office of the Permanent Secretary, Ministry of Interior (OPSMOAC), “Rayong agricultural production,” https://www.opsmoac.go.th/rayong-dwl-files-461691791909, 2024, accessed: Jan. 10, 2026.

[6] European Space Agency, “Sentinel-2 data collection,” accessed: Nov. 21, 2025. [Online]. Available: https://dataspace.copernicus.eu/data-collections/copernicus-sentinel-data/sentinel-2

[7] W. Li, H. Jiang, W. Huang, G. Yang, D. Zhang, and C. Wang, “Sentinel-2 for crop mapping and monitoring: A review,” *Journal of Imaging*, vol. 8, no. 12, p. 316, 2022.

[8] E. Commission, “Sentinel-2 for agriculture,” Copernicus — The European Union’s Earth Observation Programme, 2015, https://www.copernicus.eu/en/sentinel-2-agriculture.

[9] Copernicus Data Space Ecosystem, “Sentinel-2 level-2a collection,” 2026, accessed: Jan. 7, 2026. [Online]. Available: https://browser.stac.dataspace.copernicus.eu/collections/sentinel-2-l2a

[10] J. W. Rouse, R. H. Haas, J. A. Schell, and D. W. Deering, “Monitoring vegetation systems in the great plains with erts,” in *Proc. 3rd Earth Resources Technology Satellite-1 Symposium*. NASA Goddard Space Flight Center, 1974. [Online]. Available: https://ntrs.nasa.gov/citations/19740022614

[11] 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 Sensing of Environment*, vol. 83, no. 1–2, pp. 195–213, 2002.

[12] J. Dash and P. J. Curran, “The meris terrestrial chlorophyll index,” *International Journal of Remote Sensing*, vol. 25, no. 23, pp. 5403–5413, 2004.

[13] Z. Mondschein, A. Paliwal, T. S. Shiferaw, J. Chamberlin, R. Wang, and M. Jain, “Mapping field-level maize yields in ethiopian smallholder systems using sentinel-2 imagery,” *Remote Sensing*, vol. 16, no. 18, p. 3451, 2024.

[14] S. P. Serbin, W. K. Gower, J. W. Norman, and S. A. Townsend, “Leaf optical properties reflect variation in photosynthetic metabolism and its sensitivity to temperature,” *Journal of Experimental Botany*, vol. 63, pp. 489–502, 2012.

[15] S. Feng, J. Zhao, T. Liu, H. Zhang, Z. Zhang, and X.-y. Guo, “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*, vol. 12, no. 9, pp. 3295–3306, 2019.

[16] S. K. McFeeters, “The use of the normalized difference water index (ndwi) in the delineation of open water features,” *International Journal of Remote Sensing*, vol. 17, no. 7, pp. 1425–1432, 1996.

[17] W. Han, X. Zhang, Y. Wang, L. Wang, X. Huang, J. Li, S. Wang, W. Chen, X. Li, R. Feng, R. Fan, X. Zhang, and Y. Wang, “A survey of machine learning and deep learning in remote sensing of geological environment: Challenges, advances, and opportunities,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 202, pp. 87–113, 2023.

[18] Z. M. Mobarakeh, S. Pourmanafi, and M. Ahmadi, “Employing sentinel-2 time-series and noisy data quality control to enhance crop classification in arid environments: A comparison of machine learning and deep learning methods,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 142, p. 104678, 2025.

[19] M. Belgiu and L. Drăguţ, “Random forest in remote sensing: A review of applications and future directions,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 114, pp. 24–31, 2016.

[20] Provincial Labour Office Rayong, Ministry of Labour, Thailand, “Province map of rayong,” 2016. [Online]. Available: https://rayong.mol.go.th/en/overall/province-map

[21] D. Runfola, A. Anderson, H. Baier *et al.*, “geoboundaries: A global database of political administrative boundaries,” *PLOS ONE*, vol. 15, no. 4, p. e0231866, 2020.

[22] B.-C. Gao, “Ndwi—a normalized difference water index for remote sensing of vegetation liquid water from space,” *Remote Sensing of Environment*, vol. 58, pp. 257–266, 1996.