SPATIAL MACHINE LEARNING FOR MONITORING TEA LEAVES AND CROP YIELD ESTIMATION USING SENTINEL-2 IMAGERY, (A Case of Gunung Mas Plantation, Bogor)

Dini Nuraeni[1,2], Masita Dwi Mandini Manessa[1]
[1]Department of Geography, Faculty of Mathematics and Natural Sciences, University of Indonesia
[2]Geospatial Information Agency of Indonesia, Bogor, Indonesia
e-mail: dini.nuraeni@ui.ac.id / dini.nuraeni@big.go.id

Received: 30.12.2022; Revised:31.12.2022; Approved: 31.12.2022

Abstract. Indonesia's tea production and export volume have fluctuated with a downward trend in the last five years, partly due to the increasingly competitive world tea quality. Crop yield estimation is part of the management of tea plucking, affecting tea quality and quantity. The constraint in estimating crop yields requires technology that can make the process more effective and efficient. Remote sensing technology and machine learning have been widely used in precision agriculture. Recently, big data processing, especially remote sensing data, machine learning, and deep learning have been carried out using a cloud computing platform. Therefore, we propose using GeoAI, a combination of Sentinel-2A imagery, machine learning, and Google Collaboratory, to predict ready for plucking tea leaves at optimal plucking time at Gunung Mas Plantation Bogor. We used selected bands of Sentinel-2A and extracted more features (i.e., NDVI) as a training set. Then we utilized the tea block boundaries and tea plucking data to generate labels using Random Forest (RF) and Support Vector Machine (SVM). The classification results were further used to estimate the production of crop tea yield. The RF classifier is able to achieve an overall accuracy of 51% and SVM of 54%. Meanwhile, accuracy at optimally aged tea blocks is able to achieve 75.62% for RF and 52.88% for SVM. Thus, the SVM classifier is better in terms of overall accuracy. Meanwhile, the RF classifier is superior in predicting ready for plucking tea at optimally aged tea.

Keywords: *GeoAI, Sentinel-2, machine learning, crop tea yield estimation.*

1 INTRODUCTION

Indonesia is the eighth world's largest tea-producing country after China, India, Kenya, Argentina, Sri Lanka, Turkiye, and Vietnam, with a total production of 138,323 tons in 2020 (FAO 2022). During the period 2016–2020, Indonesian tea production and volume of tea exports fluctuated with a downward trend (BPS 2020). In 2020, Indonesia ranked tenth with 46,265 tons of total exports and fourteenth with 96.3 million USD of export value (BPS 2020, FAO 2022).

Low crop production of tea is caused by decreased plant performance, fertilization below the target, lack of support crop for infrastructure, limited quantity and skill of picking workers, and inadequate supervision (PTPN VIII 2019). The sales volume of tea in 2020 also decreased due to the over-supply in producing countries which are also potential buyers, and the scarcity of containers in Indonesia (PTPN VIII 2020). In addition, the competitive quality of tea from the world's tea-producing countries has resulted in low market absorption in Indonesia.

The quality and quantity of tea are influenced by the management of plucking tea (Dewi *et al.* 2019). It starts from harvest planning to post-harvest, which requires optimal preparation to increase tea productivity. Preparing work plans for harvest and post-harvest activities requires information on estimated crop yields. Crop yield estimation is an activity that estimates the potential production for the following harvest season. The estimated yield must reflect the actual production or be close to accuracy because it will affect the time, cost and resources needed (Junaedi 2020).

PTPN VIII uses a sampling method on a specific area in several tea blocks to get the average tea yield and estimate tea production each month. There are constraints in estimating tea yields, such as the time required for selecting sampling blocks and collecting ready-to-harvest tea samples, the size of the plantation area, the weather, and the limited number of pickers (Ricky 2022).

Meanwhile, the quality of tea plucking is influenced by the time of plucking (Setyamidjaja 2000). The right plucking time with the right plucking round and plucking area will result in optimal tea production. Some tea blocks in the Gunung Mas plantation are plucked beyond their plucking round due to a lack of pickers and the plantation foreman's visual selection of plucking areas. As a result, plucking old shoots yields a large quantity of tea but of poor quality (Tri 2022).

Technology as a supporting element in the tea value chain can increase productivity, effectiveness, and efficiency to remain competitive in the global market. Remote sensing technology has the potential to be applied in almost every aspect of agricultural precision, ranging from land preparation to harvest (Sishodia *et al.* 2020). Multispectral and multitemporal remote sensing with high spatial resolution and low cost have an advantage in collecting broad and accurate information (Sishodia *et al.* 2020). Data processing techniques such as big data analysis, artificial intelligence, machine learning, and cloud computing systems have been used to store, process, and utilize such a large amount of data for applications in agricultural precision. (Kamilaris *et al.* 2017, Zhou *et al.* 2016, Khattab *et al.* 2016, Pavo'n-Pulido *et al.* 2017).

The use of cloud computing can overcome the limitations of hardware and software bottlenecks as well as the complexity of device configuration for big data processing. Google Earth Engine and Google Collaboratory are two cloud-based platforms that provide a GPU for geospatial data analysis, particularly raster data, as well as machine learning and deep learning libraries. The combination of spatial science, artificial intelligence methods, data mining, and high-performance computing to extract information from big spatial data is known as geospatial artificial intelligence (GeoAI) (Gomez *et al.* 2016, Li *et al.* 2016, VoPham *et al.* 2018). GeoAI opens up more possibilities for using earth observation data collected from various constellations of satellites and sensors with high spatial, spectral, and temporal resolution (Pereira 2020).

The traditional linear regression model (TLRM) and normalized difference vegetation index (NDVI) are used in most empirical yield prediction performances. Still, machine learning is also an efficient empirical method for classification and prediction. Nurmalasari *et al.* (2017) used a spectral approach based on Sentinel-2 imagery with the vegetation index (NDVI, SAVI, and ARVI) and linear regression to estimate the production of tea shoots. Phan *et al.* (2020) used MODIS NDVI data to monitor tea status and forecast results using a random forest (RF) algorithm, support vector machine (SVM), and the TLRM. Kim and Lee (2016) estimated corn yields in the State of Iowa using four machine learning approaches such as RF, SVM, ERT (highly random tree), and DL (deep learning). Cai *et al.* (2019) built various empirical models for yield prediction using three mainstream machine learning methods (SVM, RF, and neural network). Empirically, the RF algorithm was regarded as a robust tool for predicting crop yield.

This study examines machine learning using RF and SVM algorithms to predict ready for plucking tea leaves at optimal plucking time using Sentinel-2 imagery. Then calculate the estimated crop yield from the prediction results.

2 MATERIALS AND METHODOLOGY

2.1 Location

The study area is located at Gunung Mas Plantation in Bogor Regency, West Java Province, Indonesia. The Gunung Mas Plantation is one of the plantations owned by PT Perkebunan Nusantara VIII (PTPN VIII) and consists of three afdeling (sectors), namely Gunung Mas 1, Gunung Mas 2, and Cikopo Selatan. The study focused on afdeling Gunung Mas 1 in Cisarua District, Tugu Selatan Village, with an area of approximately 139.25 Ha (Figure 2-1).

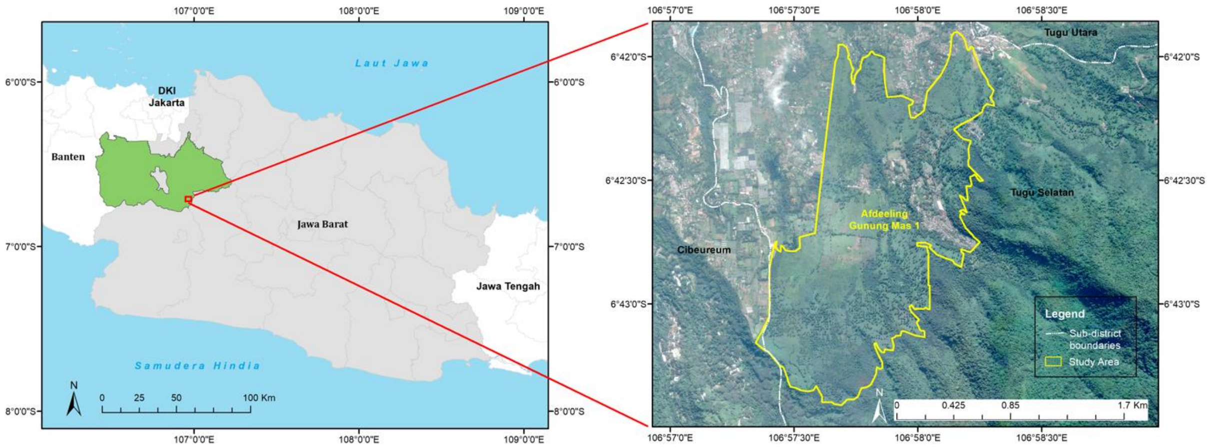

Figure 2-1: Study area location in Jawa Barat Province, Indonesia.

2.2 Data

The several data and materials used in this study as shown in Table 2-1. We used the Avenza Maps, Global Mapper, ArcGIS Pro, and Google Colaboratory platforms to collect and process data.

Table 2-1: Data used in this study

Data	Source	Specification	Date
Sentinel-2	Copernicus Open Access Hub (https://scihub.copernicus.eu/dhus/#/home)	13 spectral bands: four bands (B2, B3, B4, B8) at 10 m; six bands (B5, B6, B7, B8A, B11, B12) at 20 m; three bands (B1, B9, B10) at 60 m spatial resolution	January 2020 to July 2022
NDVI	Sentinel 2A	NIR band and Red band of Sentinel 2A	January 2020 to July 2022
Tea plantation map	PTPN VIII (Gunung Mas Tea Plantation)	Afdeling and blocks boundaries, pdf	-
Tea plucking data	PTPN VIII	Daily; tea age; plucking schedule; crop yield data per day per blocks	January 2020 to July 2022
Crop yield estimated data	PTPN VIII	Monthly per blocks	January 2020 to July 2022
Basemap	Geospatial Information Agency (BIG)	Very high-resolution imagery (Pleiades & Worldview, spatial resolution 0,5 m); administrative boundary	2017; 2019
Field survey data	GPS (Avenza Maps), interview	Longitude, latitude; plantation manager or plantation supervisor	August 2022

2.2.1 Sentinel-2

We utilized time series remote sensing satellite imagery as input in predicted ready for plucking tea leaves to estimate tea yields. The launch of Sentinel-2 in June 2015 has overcome many limitations in applying remote sensing techniques for precision agriculture in previous years (Segarra et al., 2020). Sentinel-2 imagery is one of the multitemporal and multispectral remote sensing satellite data. It has a resolution of 10 meters and a wide range. This study used cloud-free Sentinel 2A imagery in the research area from January 2020 to July 2022. We obtained 11 scenes of Sentinel-2A imagery, with eight scenes as training data and three as testing data.

Table 2-2: Data used in this study.

Year	Month											
	1	2	3	4	5	6	7	8	9	10	11	12
2020							1	1			1	
2021			1	1	1		1	1				
2022			1				2					

2.2.2 Tea Plucking Data

The data on tea plucking at the Gunung Mas plantation contains information on block area, age of tea (counted after the last plucking), age of pruning, and the number of crop yields per day for each block. Ideally, the tea is plucked after the tea age enters the plucking round, which is set for each block because the delay in plucking will affect the quality and productivity of the tea produced. The plucking round in the Gunung Mas plantation is determined by the tools used. Scissor-plucked tea blocks have a plucking round of 25—30 days, while machine-plucked tea blocks have a 45—50 day plucking round. (Tri 2022). Based on empirical experience in the Gunung Mas plantation, tea leaves will be aged on the 35th day (for tea plucked using scissors) and the 60th day (for tea plucked using machines). So, the optimal plucking time for scissors-plucked block tea is 25—34 days and 45—59 days for machine-plucked block tea.

2.2.3 The Other Data

This study also used base map data, a map of tea plantations, estimated crop yield data, and field survey data. The base map data is in the form of very high-resolution imagery and administrative boundaries. The data is used as a guide for digitizing the Gunung Mas plantation map and verifying land use/land cover (LULC). Information regarding the boundaries of afdelings, blocks, and other LULCs is obtained from the tea plantation map. The estimated crop yield is the monthly data per afdeling from PTPN VIII, which was obtained by sampling several tea blocks in an area of 10 m x 10 m to determine the average crop yield. The field survey data were collected through ground checks and interviews with plantation managers or supervisors. The survey was conducted in August 2022 using the Avenza Maps application to take coordinates and photos, especially of LULC changes. The data is further used for LULC validation and classification.

2.3 Methods

In this study, Sentinel-2 was composited into five bands (blue, green, red, NIR, and SWIR). Then we generate labels using block boundaries, field surveys, and tea-plucking data. Two popular machine learning algorithms were used to forecast ready for plucking tea leaves at the optimal time: random forest (RF) and support vector machine (SVM). The prediction results were further used to estimate crop yield production (Figure 2-2).

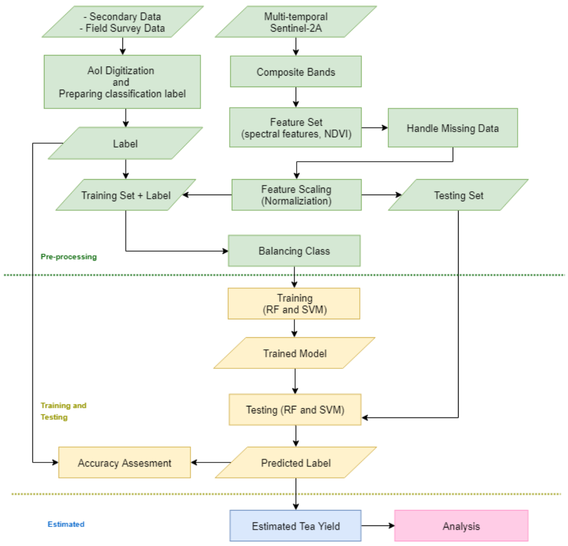

Figure 2-2: Research workflow.

2.3.1 Pre-Processing Data

Since this study uses supervised learning, preparing a labeled training set is necessary. We used ground check data to verify secondary data for tea blocks to obtain the most recent conditions in the Gunung Mas plantation. Then we dug for information from plantation managers and supervisors regarding plucking round, tea age, picking methods, and tea production. The results were further used to generate labels.

Labels are created using only productive tea blocks; tea blocks that have changed land use are also excluded (Figure 2-3). The tea is then divided into two classifications: ready for plucking (1) and not ready for plucking (0). Tea plants with an optimal age are classified as ready for plucking, whereas tea plants that have not yet entered their optimal age or have passed the optimal age are classified as not ready for plucking. The next step is to resize the composited Sentinel-2 data to equalize the image size. In this study, the image was resized to 32x32 pixels. Resize is also intended to speed up the image recognition process by machines. All the preparation steps for the labeled training set above were carried out using ArcGIS Pro software.

We used Google Colaboratory and the Python programming language for further data processing. The feature set in the form of an image must be converted into another form that the machine can learn before being processed using machine learning. So we converted the feature set into arrays for the machine to learn. We also used Sentinel-2 NDVI as the feature set since NDVI is often used as a supporting tool for accurate and efficient yield estimation for some crops such as corn, rice, tea, and wheat. (Kim and Lee 2016, Franch *et al.* 2021). The NDVI algorithm is as follows:

$$NDVI = \frac{NIR - Red}{NIR + Red} \ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

N = Band Near-Infrared Radiation, RED = Band Red.

Missing data is relatively common in almost all research and can significantly impact trial validity and lead to invalid conclusions (Graham 2019, Kang 2013). Therefore, for this study, we used listwise deletion as the most common method of handling missing data in the feature set. Then, to ensure the training set has the same scale, it is necessary to do feature scaling. The feature scaling method is used to normalize the range of independent variables or data features and assist them all in the same range. We scale the feature set using the normalization method, so the values are between the range of [0.1] or [-1.1]. Normalization of data will help speed up the learning process in machine learning (Li & Liu, 2011). Then we divide the data into the training and testing sets with a ratio of 80:20. The general formula for normalization is as follows:

$$Xnorm = \frac{X - min(X)}{(X) \; - (X)} \ldots\ldots\ldots\ldots (2)$$

X = value from the feature X
min(X) = the feature's minimum value
max(X) = the feature's maximum value

In this study, we have an unbalanced portion of the training set between one class and another, this condition is known as a class imbalance (Figure 4). When there is a class imbalance, the minority class is more challenging to predict than the majority class. However, there are times when the minority class has more important information, which can affect the accuracy (Windyaning and Suprapto 2020). Therefore, we used random undersampling methods to balance our training set. It means that we randomly remove majority class examples from the training set (Weiss 2013) and only use as many as the minority class has.

2.3.2 Predicting Ready for Plucking Tea

Two machine learning algorithms, random forest (RF) and support vector machine (SVM), were used to forecast ready for plucking tea at the optimal plucking time. RF is an ensemble algorithm capable of performing regression and classification tasks using multiple decision trees as base learning models and a technique known as bootstrap and aggregation (Breiman 2001, Zhu *et al.* 2019). The decision tree classifies a data sample whose class is unknown into existing classes. RF selects a random sample from the training set, creates a decision tree, and gets a prediction. It then performs a vote for each prediction and takes the result with the majority of votes in the case of classification or the average in the case of regression (Ali *et al.* 2012). This operation repeats for the assigned number of trees.

SVM is a machine learning algorithm for classification, linear and non-linear regression analysis, and time series prediction. SVM determines the best hyperplane to separate data sets from two or more different classes (Cortes 1995). It has advantages such as maximizing the margin between classes using a support vector, which speeds up the computing process (Cortes 1995, Gandhi 2018). The ready for plucking tea leaves was predicted from the available data, such as NDVI, tea ages, plucking round, and historic tea crop yield.

2.3.3 Estimated Tea Crop Yield

The prediction results were further used to estimate crop yields in blocks with optimal plucking time. The pixel area with the ready for plucking classification results is calculated and compared with the block area with the optimal plucking time. Then the estimated yields of the two were compared. Estimated crop yield using the following formula:

$$Tp = \left(\frac{La}{Ls}\right) * Hs \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3)$$

Tp = crop yield estimated
La/Ls = crop area
Hs = average crop yield

3 RESULTS AND DISCUSSION

The LULC in the Gunung Mas plantation is divided into six classifications: built-up area, agro-tourism, tea, forest, cinnamon, and river/gap. Figure 3-1 (a) shows the LULC and distribution of ground checks carried out to obtain the current condition of LULC in the Gunung Mas plantation. However, the problem of access to several blocks makes ground checks uneven. The ground check found wasteland, avocado, shrubs, and mixed vegetation in the Gunung Mas plantation area. Thus, there was a total change in LULC in three tea blocks and a partial change in LULC in six tea blocks, resulting in a reduction in the tea area (Figure 3-1 (b)). Insignificant vegetation was ignored in the recent delineation of the tea block boundary. The result is 21 of 23 tea blocks with updated boundaries as input for the training set (Figure 3-1 (c)).

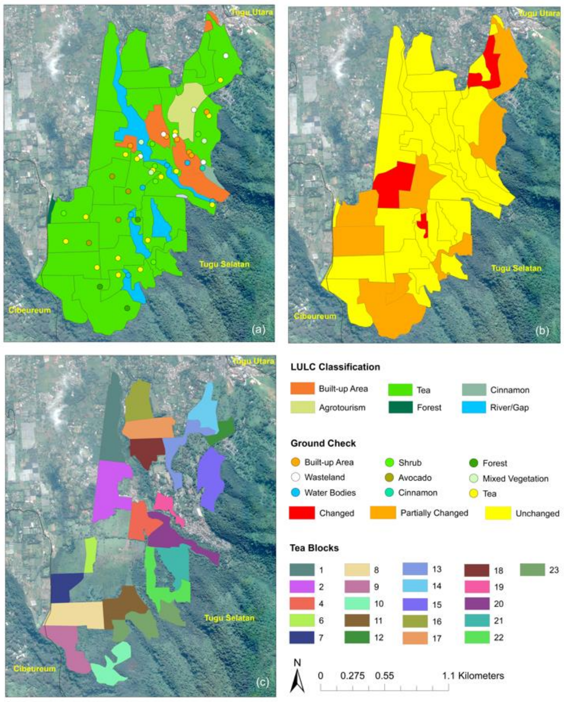

Figure 3-1: LULC change and productive tea block for training set labels.

The training set label is based on the optimal age of the tea to be plucked. The optimal plucking time for scissors-plucked block tea is between 25—34 days and 45—59 days for machine-plucked block tea. However, from tea plucking data and field observations, it is known that some of the plucked blocks do not follow the picking round or the optimal time. Figure 3-2 shows three blocks of tea that were ready for plucking but not yet harvested on August 27, 2020. Meanwhile, the four harvested tea blocks had already passed their optimal age. Then, on April 24, 2021, one of the four optimal age blocks was harvested. The other three blocks are harvested past their optimal age.

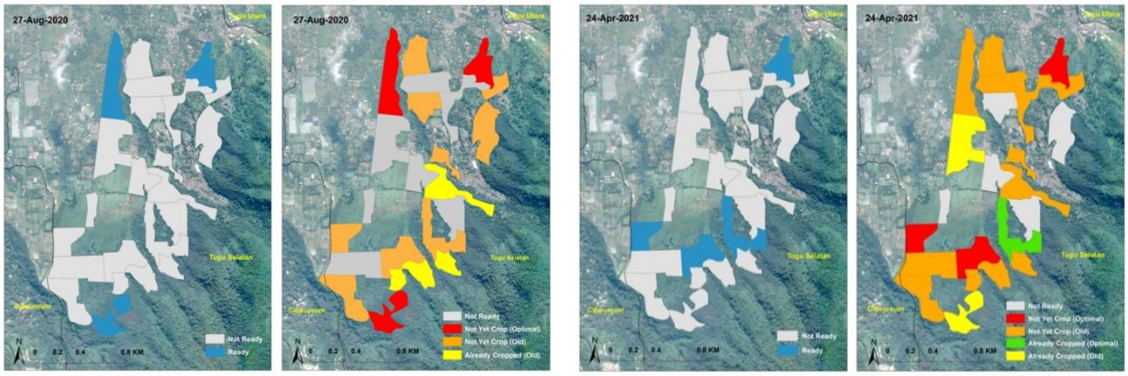

Figure 3-2: The suitability of the actual tea plucking time with the optimal age of tea

Training labels are classified into two types: “ready” and “not ready”. Tea leaves that have reached their optimal age will be included in the “ready” category with a value of 1. Meanwhile, those who have not yet entered or have passed the optimal age are classified as “not ready” with a value of 0. The classification distribution of eight Sentinel-2 images used as a training set is shown in Figure 3-3.

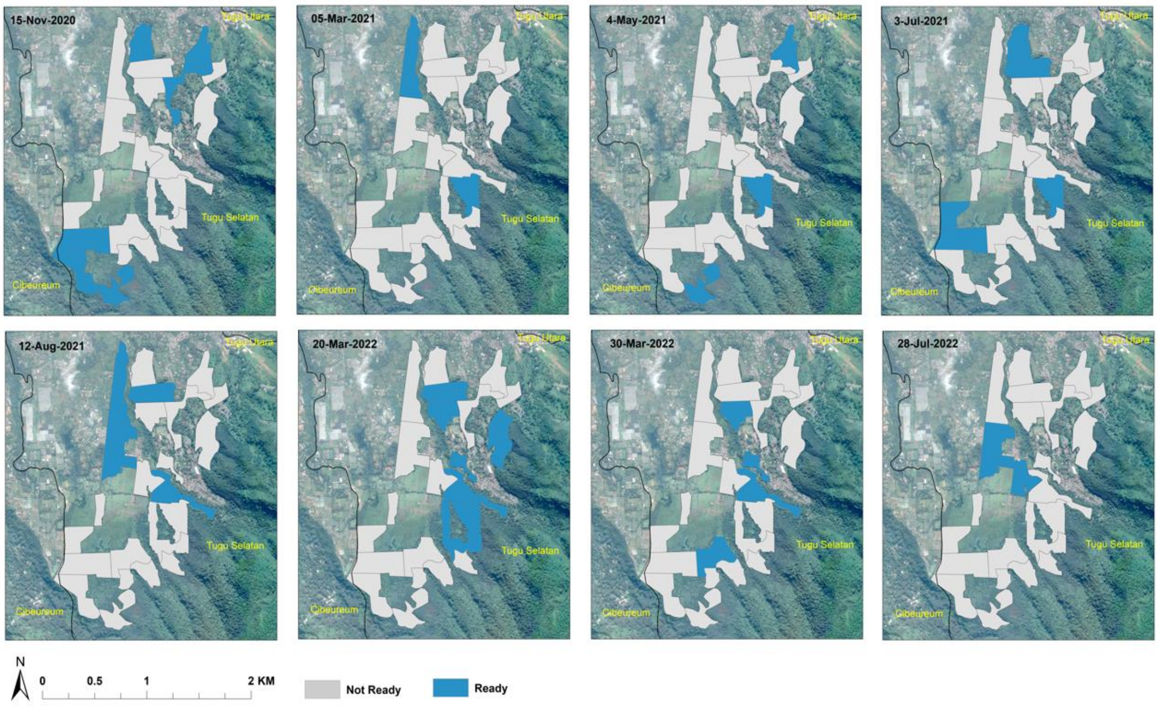

Figure 3-3: The classification distribution of training set before balancing class.

There is a class imbalance between the "ready" category as the minority class and the "not ready" category as the majority class. Before class balancing, the number of pixels with a value of 0 is 11,580,080, and the number of pixels with a value of 1 is 95,296. Due to the random undersampling method that we use to balance our training set, pixels with a value of 0 are deleted randomly until they reach the same number of pixels with a value of 1. As a result, each class has 95,296 pixels.

The prediction results of ready for plucking tea leaves at optimal plucking time in the testing set are shown in Figure 3-4. Red pixels indicate tea leaves that are not ready to be picked, while green pixels indicate tea leaves that are ready to be picked. As a result, overall accuracy with the RF algorithm is able to achieve 51% and with the SVM 54%. Meanwhile, accuracy at optimally aged tea blocks is able to achieve 75.62% for RF and 52.88% for SVM (Table 3-1). Furthermore, the results were further used to estimate crop yield production.

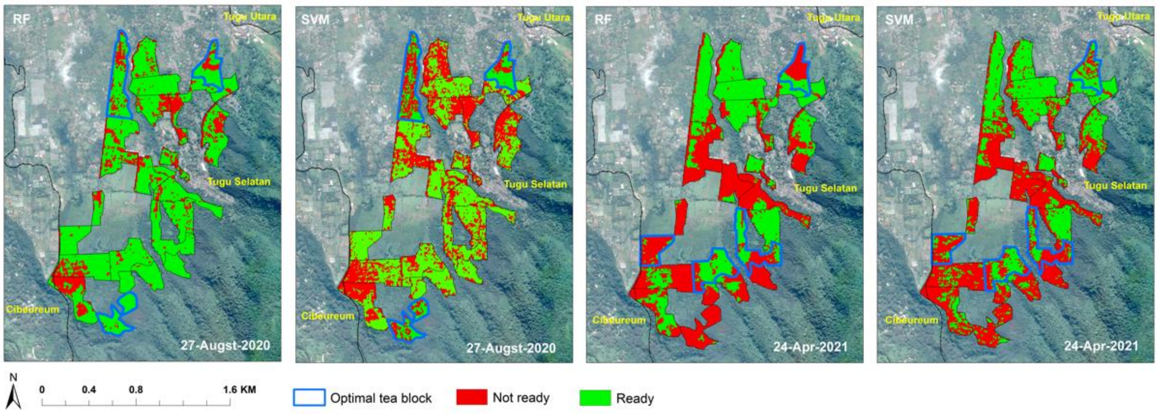

Figure 3-4: Result of ready for plucking tea leaves.

Table 3-1: Result of accuracy.

Date	Overall Accuracy (%)		Accuracy at Optimal Tea Block (%)	
	RF	SVM	RF	SVM
27 August 2020	30	44	75.62	52.88
24 April 2021	51	54	45.04	45.45

Table 3-2: Average crop yield per month.

Month	Crop Yield Production (kg)	Crop Area (ha)	Average Crop Yield (kg/ha)
August 2020	40,160	81.99	489.82
April 2021	64,420	76.61	843.49

We compare the estimated crop yield using the predicted optimal block area to the actual optimal block area. The average crop yield in one month was obtained using a formula based (3) on total crop yields and tea plantation area (Table 3-2).

As a result of the two testing sets, the estimation of crop tea yield on tea blocks of optimal age is able to achieve 72.47% for RF algorithms and 50.68% for SVM (Table 3-3).

Table 3-3: Comparison of crop yield estimation results between the actual area and the predicted area.

Date	Crop Yield Area (ha)			Crop Yield Production (kg)			Percentage of predicted production to actual (%)	
	Actual	RF	SVM	Actual	RF	SVM	RF	SVM
27 August 2020	22.81	16.53	11.56	11,173	8,087	5,662	72.47	50.68
24 April 2021	29.40	11.89	12.00	24,799	10,029	10,122	40.44	40.82

4 CONCLUSION

Good tea plucking management and proper plucking time are required for optimal tea production. The fact that some of the tea blocks in the Gunung Mas plantation were harvested when they were past their optimal age can have an impact on the quality of the tea produced. Estimating tea crop yields as part of planning for future harvest and post-harvest activities takes significant time and effort. The result of the overall accuracy of estimated tea crop yields with the RF classifier is able to achieve 51% and 54% with the SVM. Meanwhile, the accuracy of optimally aged tea blocks with the RF classifier is able to achieve 75.62% and 52.88% with the SVM. Thus, of the two algorithms used, the SVM classifier is better in terms of overall accuracy. Meanwhile, the RF classifier is superior in predicting ready for plucking tea leaves at optimal plucking time. So the results obtained by the RF classifier have estimated crop tea yields that are close to the results of the field data of 72.47%.

Although Sentinel-2A has limitations in providing multi-temporal datasets in mountainous areas due to clouds and rain, the results obtained in this study are sufficient for estimating crop tea yields at Gunung Mas Plantation. Further research could combine multi-source remote sensing data and involve parameters affecting shoot growth and tea production.

ACKNOWLEDGEMENTS

This research was supported by the Department of Geography, University of Indonesia, and the National Research and Innovation Agency through the continuation scholarship program. We would like to thank PTPN VIII Gunung Mas Plantation for providing data and information and permitting research in the area. We are also grateful to the reviewers and editors for their careful reading which helps us prepare a more improved manuscript.

REFERENCES

Ali J., Khan R., Ahmad N., Maqsood I., (2012), Random forests and decision trees. International Journal of Computer Science Issues 9(5):272-278

Badan Pusat Statistik (2020), Statistik Teh Indonesia (Katalog/Catalog:5504001). BPS RI/BPS-Statistics Indonesia, Jakarta.

Breiman, L., (2001), Random forests. Mach Learn 45:5-32. doi:10.1023/A:1010933404324

Cai Y., Guan K., David Lobell, et al, (2019), Integrating Satellite And Climate Data To Predict Wheat Yield In Australia Using Machine Learning Approaches. Agricultural and Forest Meteorology 274:144–159. doi: 10.1016/j.agrformet.2019.03.010

Cortes C., Vapnik V., (1995), Support Vector Networks. Mach Learn 20:273-297. doi:10.1007/BF00994018

Dewi Synthia, Purnomo, (2019), Mutu Petik Teh (Camellia sinensis (L.) O. Kuntze) di Kebun Bedakah, Wonosobo, Jawa Tengah. Buletin Agrohorti 7(1):337-342.

Food and Agriculture Organization of the United Nations (FAO), FAOSTAT. Available via https://www.fao.org/faostat/en/#rankings/countries_by_commodity, https://www.fao.org/faostat/en/#rankings/countries_by_commodity_exports. Accessed 14 March 2022

Franch Belen, Bautista Alberto San, Fita David, et al, (2021), Within-Field Rice Yield Estimation Based on Sentinel-2 Satellite Data. Remote Sens 13:4095. doi:10.3390/rs13204095

Gandhi Rohith, (2018), Support Vector Machine - Introduction to Machine Learning Algorithms. Available via https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47. Accessed 12 Apr 2022.

Gomez C., White Joanne C., Wulder Michael A., (2016), Optical remotely sensed time series data for land cover classification: A review. ISPRS Int J Geo-Info 116:55-72. doi:10.1016/j.isprsjprs.2016.03.008

Graham JW., (2009), Missing data analysis: making it work in the real world. Annu Rev. Psychol 60:549–576. doi:10.1146/annurev.psych.58.110405.085530

Junaedi, Thamrin Syahruni, Syafaat Muhammad, et al, (2020), Taksasi Produksi Tanaman Kopi dengan Metode ABC. Jurnal Agroplantae 9(2): 66-75.

Kang H, (2013), The prevention and handling of the missing data. Korean J Anesthesiol 64(5):402-406. doi: 10.4097/kjae.2013.64.5.402

Kamilaris A., Kartakoullis A., Prenafeta-Boldú F.X., (2017), A review on the practice of big data analysis in agriculture. Comput. Electron. Agric 143:23–37. doi:10.1016/j.compag.2017.09.037

Khattab A., Abdelgawad A., Yelmarthi K. (2016), Design and implementation of a cloud-based IoT scheme for precision agriculture. Paper presented at the 28th International Conference on Microelectronics (ICM). Giza, Egypt. 17-20 Dec 2016

Kim N., Lee Y.W., (2016), Machine Learning Approaches To Corn Yield Estimation Using Satellite Images And

Climate Data: A Case Of Iowa State. J. Korean Soc. Surv. Geod. Photogramm. Cart 34(4):383–390. doi:10.7848/ksgpc.2016.34.4.383

Li Songnian, Dragicevic S., Castro F.A., et al, (2016), Geospatial big data handling theory and methods: A review and research challenges. ISPRS Int J Geo-Info 115:119-133. doi:10.1016/j.isprsjprs.2015.10.012

Li W., Liu Z., (2011), A method of SVM with normalization in intrusion detection. Procedia Environ, 11:256-262. doi:10.1016/j.proenv.2011.12.040

Nurmalasari Intansania, Santosa Sigit H.M.B., (2018), Pemanfaatan Citra Sentinel-2A untuk Estimasi Produksi Pucuk Teh di Sebagian Kabupaten Karanganyar. Jurnal Bumi Indonesia 7(1).

Pavo'n-Pulido N., Lo'pez-Riquelme J.A., Torres R., et al, (2017), New trends in precision agriculture: A novel cloud-based system for enabling data storage and agricultural task planning and automation. Precision Agric 18:1038–1068. doi:10.1007/s11119-017-9532-7

Pereira A.G., Ojo Adegboyega, Curry Edward, et al, (2020). Data Acquisition and Processing for GeoAI Models to Support Sustainable Agricultural Practices. Paper presented at the 53rd Hawaii International Conference on System Sciences. Maui, Hawai. 7-10 Jan 2020

Phan Phamchimai, Chen Nengcheng, Xu Lei, et al, (2020), Using Multi-Temporal MODIS NDVI Data to Monitor Tea Status and Forecast Yield: A Case Study at Tanuyen, Laichau, Vietnam. Remote Sens 12:1814. doi:10.3390/rs12111814

PT. Perkebunan Nusantara VIII, (2019), Memperkuat Portofolio Komoditas dan Optimalisasi Aset Perusahaan. In: Laporan Tahunan 2019. Available via https://www.portalpn8.net/Content/Images/AnualReport/Annual%20Report%202019/AR%20PTPN%202019%20SPREAD%2016%20OKT%20-%20Judul%20super%20fix.pdf. Accesed 13 Apr 2022

PT. Perkebunan Nusantara VIII, (2020), Membangun Fondasi Menuju Keberlanjutan. In: Laporan Tahunan 2020. Avalaiable via https://www.portalpn8.net/Content/Images/AnualReport/Annual%20Report%202020/AR%20PTPN%20VIII%202020%20Rev%201.pdf. Accesed 4 Jul 2022

Ricky, (2 Aug 2022). Personal interview.

Setyamidjaja D, (2000), Teh Budidaya dan Pengolahan Pasca Panen. Kanisius, Yogyakarta.

Sishodia R.P., Ray R.L., Singh S.K., (2020), Applications of Remote Sensing in Precision Agriculture: A Review. Remote Sens 12:3136. doi:10.3390/rs12193136

Tri, (3 Aug 2022). Personal interview.

VoPham Trang, Hart Jaime E., Laden Francine, et al, (2018), Emerging trends in geospatial artificial intelligence (geoAI): potential applications for environmental epidemiology. Environmental Health 17:40. doi:10.1186/s12940-018-0386-x

Weiss G.M., (2013), Foundations of Imbalanced Learning. In: Imbalanced Learning: Foundations, Algorithms, and Applications. New Jersey: IEER press & Wiley, New Jersey, p 13-41

Windyaning U., Suprapto S., (2020), Oversampling Method To Handling Imbalanced Datasets Problem In Binary Logistic Regression Algorithm. IJCCS 14(1):1-10.

Zhou L., Chen N., Chen Z., et al, (2016), ROSCC: An efficient remote sensing observation-sharing method based on cloud computing for soil moisture mapping in precision agriculture. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 9(12):5588–5598. doi:10.1109/JSTARS.2016.2574810