



Estimated Shallot Yield Area Using the Rapid Classification of Croplands Method

Agung Budi Santoso^{1*}, Tumpal Sipahutar², Tomy Purba², Sarman Paul Lumbantobing², Shabil Hidayat², Moral Abadi Girsang², Lermansius Haloho²

(Received January 2024/Accepted December 2024)

ABSTRACT

Shallots are one of the horticultural commodities that have fluctuating prices. Market integration occurs horizontally but not vertically due to poor information systems at the producer and consumer levels. This study aimed to estimate the area of shallot land quickly using the rapid classification of croplands method. The research was conducted in Merek District, Karo Regency, North Sumatra. Primary data obtained from survey activities were processed using the Google Earth Engine platform. Classification and regression trees (CART) and random forest (RF) algorithms were used to classify land cover as onion and non-onion classes. The shallot land area based on this method was 74.4 hectares, with an area accuracy of 95% (RF) and 24% (CART) and a location accuracy of 92% (CART and RF). The rapid classification of croplands method can estimate land area quickly. It helps stakeholders who need information on shallot production projections and can be developed to improve the vertical market integration information system (market integration between producers and consumers). Some areas for improvement of this method are limited access and resolution, inability to describe up to the level of garden bunds, and the condition of the area covered by clouds, which will affect the accuracy of the results.

Keywords: shallots, production estimation, google earth engine, remote sensing

INTRODUCTION

Fluctuations in strategic food prices are believed to be one of the triggers for inflation (Firdaus 2021). Based on the law of supply and demand, prices are formed based on the balance of available quantities with requested quantities. Uncontrolled price fluctuations are closely related to unstable supply conditions (Serra and Gil 2013). As part of strategic food commodities, horticulture has characteristics different from those of major food commodities such as rice, corn, and soybeans. Shallots, garlic, and cayenne pepper commodities are commodities that have experienced an increase in price variability at a time when rice, beef, and eggs have experienced a decrease in price variability during the COVID-19 pandemic (Firdaus 2021). Price fluctuations

are closely related to the balance of supply and demand for a commodity. Horticulture has different characteristics from food and plantation commodities. Generally, horticultural products are consumed in fresh conditions, so post-harvest handling differs from food commodities or plantations. The shelf life of food crops and plantations can be extended by regulating the moisture content (Aspari *et al.* 2023; Rahayu *et al.* 2014). Meanwhile, horticultural commodities must maintain freshness or moisture content until they reach consumers. This condition makes this commodity perishable, impacting the high post-harvest handling cost (Samad 2016). In addition, the relatively short shelf life makes the price drop drastically when there is an excess of stock, and vice versa; the price increases when there is a shortage of stock.

Shallots are a commodity that has a broader range of consumers than cayenne pepper. Based on (BPS2023b), the consumption of shallots in North Sumatra reaches 0.294 kg per capita/month. In comparison, garlic reaches 0.151 kg per capita/month, cayenne pepper reaches 0,160 kg per capita/month, and red chili reaches 3.6 ounces per capita/per month. Shallot consumption ranked first nationally, with a consumption rate of 0.235 kg per capita/month. Meanwhile, garlic, cayenne pepper, and red chili were

¹ Research Center for Macroeconomics and Finance, National Research and Innovation Agency, Gedung Sasana Widya Sarwono Lantai 7, Jakarta 12710, Indonesia

² Research Center for Cooperative, Corporation, and People's Economy, National Research and Innovation Agency, Gedung Sasana Widya Sarwono Lantai 7, Jakarta 12710, Indonesia

* Corresponding Author: E-mail: agun043@brin.go.id

0.163, 0.180, and 0.170 kg per capita/month (BPS 2023a).

One of the efforts to improve the price balance is to make accurate and fast estimates of production and demand. Projected land area is one of the critical indicators in estimating production. The accuracy of land projections will also improve coordination between farmers as producers and traders as a trade route to the end user (Septiana *et al.* 2017). Supply chain performance needs to be improved through partnerships, collaboration, institutional strengthening, and the availability of information systems, especially during the off-season. This available information is needed faster to avoid mistakes in the decision on the number of orders and the arrival time of shallots from outside the region.

One of the methods that can consider the estimated area of agricultural land is remote sensing. This method has been widely used in food crop commodities to determine the estimated harvest area, rice growth phase area, and productivity based on the Normalized Difference Vegetation Index (NDVI) recorded by satellites (Marsuhandi *et al.* 2020; Musfiza *et al.* 2023; Permatasari *et al.* 2021). In addition to using an index such as the NDVI, the Rapid Classification of Croplands uses all the bands in a image to classify land cover (Flood 2013). Estimating land area in horticultural commodities using remote sensing still needs to be done because farmers cultivate horticulture not in one large expanse. Based on this description, this study aims to provide an alternative method of estimating the area of shallots using remote sensing. The research results are expected to help the government develop a strategy to stabilize the price of shallots.

because Karo Regency has the second-highest contribution to shallot production in North Sumatra after Simalungun Regency, covering 25.21% (BPS 2023c). In addition, Karo Regency is relatively closer to Medan City as the destination for shallot marketing than Simalungun Regency.

Procedure

The survey was employed to obtain primary data in the form of coordinates for shallot cultivation in Merek District in November 2023. The data consisted of two classifications: the shallot group, which was the coordinate point that represents the area/expanse of shallot cultivation, and the non-shallot group, which was the coordinate point that represents the non-shallot area (housing, forest, other crops, other than shallots, roads). The survey emphasized commodities cultivated by horticultural farmers in Merek District; the survey results were obtained from 13 non-shallot locations and 12 shallot locations. The non-shallot classification, which consists of housing, roads, and forests, was carried out directly through the Google Earth engine because it is easy to identify. The final sample consisted of 100 points and 10 polygons for class 0 (shallots) and 33 points and 50 polygons for class 1 (non-shallots).

The primary data was then processed using the Rapid Classification of Croplands (Bofana *et al.* 2020) with the help of the Google Earth Engine (GEE) platform. This study's programming command (syntax) adopted programming in the GEE community. The data processing process in this study was conducted in stages (Figure 1).

In the first stage, secondary data in the form of a map of the administrative boundaries of Merek District was included in the GEE as an Area of Interest (AOI), which functions as the limit of the algorithm calculation process. Second, satellite imagery and visuals were obtained by accessing sentinel-2 data in the GEE platform. Spectral bands were used to obtain clear images and as a classification input using medoid (a multidimensional median) (Flood 2013). Third, input of the training and

METHODS

Research Site

The study was conducted in October 2023 in Merek District, Karo Regency. The location was chosen

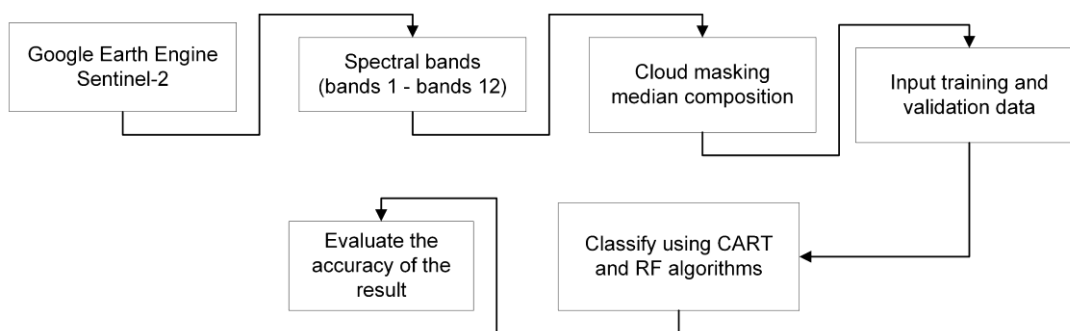


Figure 1 Land cover classification methodology (Loukika *et al.* 2021).

validation data in the form of coordinate points and accompanying information, information 0 for the shallot area, and 1 for the non-shallot area. Fourth, classification used 2 algorithms: classification and regression trees (CART) and random forest (RF). CART is a binary classification tree that provides decisions based on logical scenarios, while RF includes multiple decision trees to determine classifications (Loukika *et al.* 2021). Fifth, calculated the area of classification results.

The results from the GEE process are then verified in the field to determine the level of accuracy by comparing the yield of the shallot land area with information from field extension workers. The level of accuracy was obtained by comparing the difference between the prediction of the land area of GEE yields and the report on the area of shallots of agricultural extension workers (Marsuhandi *et al.* 2020).

$$\text{Area accuracy} = \left(1 - \frac{|\text{Area as predicted by GEE} - \text{Area according to the extension workers}|}{\text{Area according to the extension workers}}\right) \times 100 \quad (1)$$

Another accuracy was determined by verification of the GEE point with the actual plant conditions in the field (Mulyaqin *et al.* 2022):

$$\text{Area accuracy} = \frac{\text{Total of the true verified location}}{\text{Total of the verified location}} \times 100\% \dots \dots \dots (2)$$

harvest area in Karo Regency, which amounted to 943 hectares. Not only shallots, but Merek District also contributes to the production of other vegetables such as curly chili (13%), cayenne pepper (12%), potatoes (15%), cabbage (15%), napa cabbage (10%), tomatoes (17%), and carrots (5%).

Google Earth Engine (GEE) is a cloud-based geospatial platform that conducts environmental analysis and monitoring on a large scale (Figure 2). It can be accessed for free with the support of machine learning and processors owned by Google. GEEs can pull image archives from satellites, process them, and help analyze them by merging the bands they belong to (Tamiminia *et al.* 2020).

Data Training and Validation as a Field Data Input Process

Primary data in coordinate points of shallot and non-shallot planting locations were included in the GEE as input variables. One location can produce 5 to 10 training points and 1 polygon. This was done at shallot planting locations with a class 0 code (green, Figure 3A) and non-shallot locations with a class 1 code (yellow, Figure 3B). The training and validation data input variables were then processed in machine learning for class identification based on the index generated by merging bands in satellite images.

RESULTS AND DISCUSSION

Merek District is at an altitude of 920–16.20 m above sea level and has an area of 125 km², contributing the most to the production of shallots in Karo Regency. Based on data from (BPS Karo 2023), the shallot harvest area in 2022 was 412 hectares, 43% of the total shallot

Satellite Imagery and Land Cover Classification

This study used images from the sentinel-2 satellite from October 2, 2023, to November 2, 2023, adjusting the data collection period in the field. The results of the cloud masking process from the sentinel-2 image set produced a mosaic image (Figure 4B).

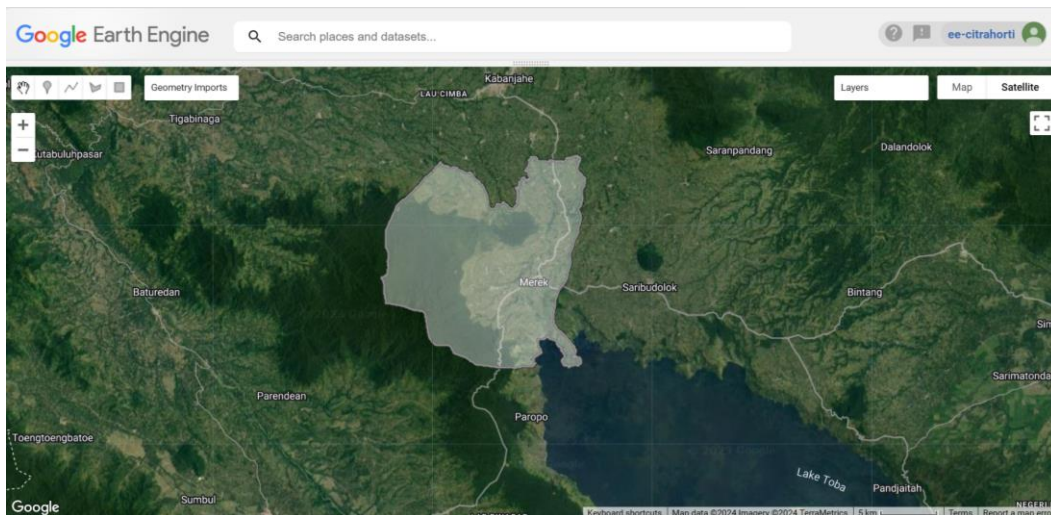


Figure 2 The location of the research (Merek District, Karo Regency) in the Google Earth Engine view.

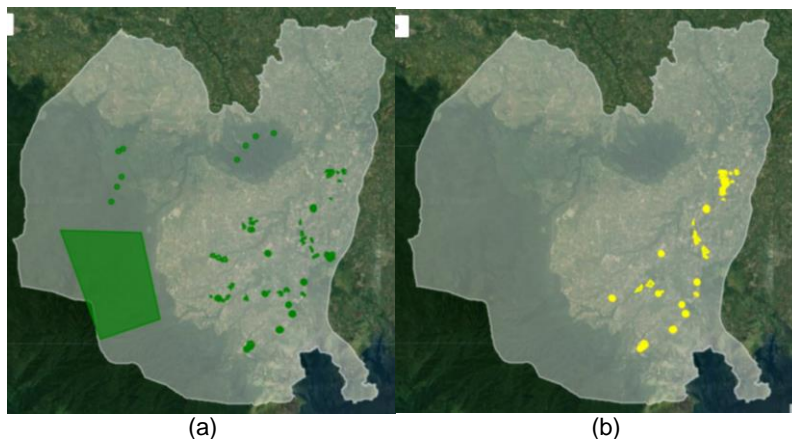


Figure 3 Data training (A) and data validation (B) based on field surveys.

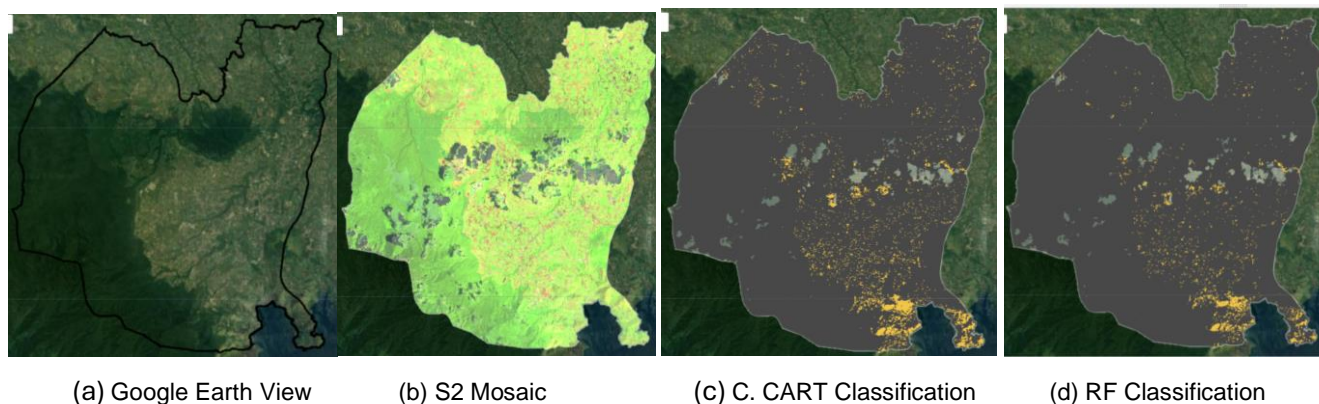


Figure 4 AOI display on GEE (A), sentinel-2 image processing results (B), CART algorithm classification results (C), and RF algorithm classification results (D).

Based on the bands in the sentinel-2 image, the classification was carried out using the CART algorithm (Figure 4C) and the RF algorithm (Figure 4D). Black indicates the non-onion classification (class 1), while orange indicates the onion classification (class 0). The accuracy validation generated by CART was 72.13%, while RF produced an accuracy validation of 71.33%. These values are slightly lower than the Phan *et al.* (2020) report, with a validation of 84.31% (RF). Phan *et al.* (2020) employed a more general classification with the division of agricultural classes, burned areas, open land, grasslands, mixed grasslands, settlements, forests, and water, while this study focuses on the division of specific commodity classes of shallots to non-shallot classes that have the potential to have the same characteristics in one agricultural class. The results of this study are closer to those of Zulkarnain and Marsisno (2022), who also estimate the land area of shallot commodities with an accuracy of 78.6% (CART).

Test of Accuracy of Estimating the Area of Shallots Using GEE

The accuracy test was carried out to evaluate the output of the estimated area of GEE shallots. This accuracy test consists of 2 stages: area accuracy and location accuracy. Area accuracy results in better accuracy obtained by RF than CART classification. Based on the extension workers' report, the area of shallot land in October 2023 was 71 hectares. Meanwhile, the estimated land area from the RF classification was 74.4 hectares or a difference of 3.4 hectares (95% accuracy). Meanwhile, the CART classification produced an estimated 125 hectares of land area, with an accuracy of only 24%.

The area accuracy decreased in November. The area of shallot land was 80 hectares. The RF classification estimated the area of shallot land to be 99.2 hectares (76% accuracy). The accuracy of the RF classification is still better than that of the CART classification, which estimates the area of shallots to be 155 hectares (6% accuracy). The CART classification in Figure 4C shows

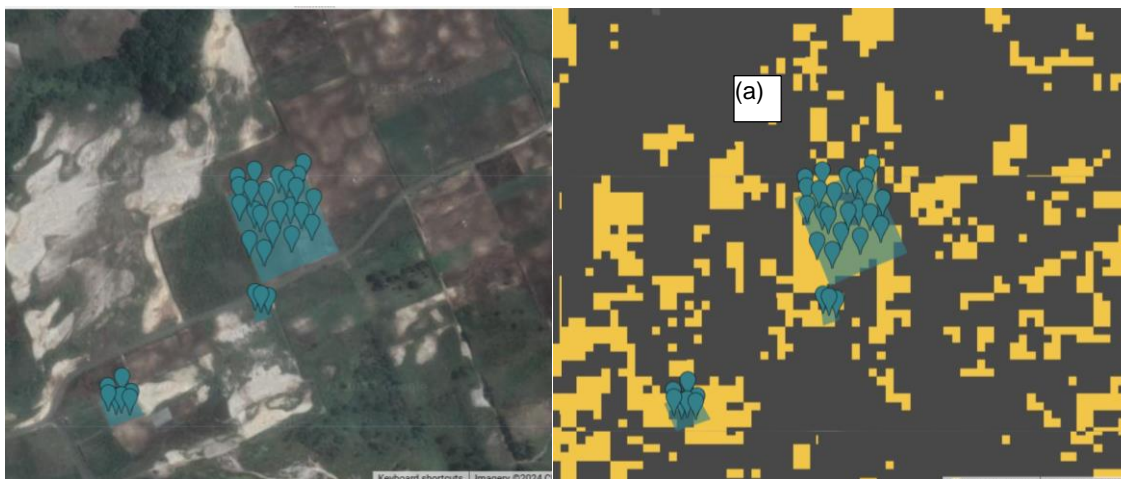
that the shades and orange dots have a wider distribution than the RF classification (Figure 4D), showing that the RF is better than the CART classification.

Based on the location accuracy test, the land area estimation method using the Rapid Classification of Croplands method produced an accuracy of 92% for both CART and RF. The exact location points mapped by GEE are 23 locations out of a total of 25 locations, as shown

in Table 1. The two points the GEE wrongly predicted were cloud-covered areas, so they were not identified. It is a weakness of techniques that rely on image sources from satellite archives, especially in foggy highland areas. In addition, weather conditions will also affect the level of accuracy; the rainy season in many areas is covered with clouds. This weakness can be minimized by widening the time range of sentinel-2 datasets or using a

Table 1 Accuracy test of GEE prediction results with actual conditions (ground checking November 2023).

Latitude	Longitude	Remarks	RF	Conclusion	CART	Conclusion
3.012312°	98.537987°	Cabbage	1	True	1	True
2.919230°	98.486790°	Shallot	0	True	0	True
2.912340°	98.519500°	Shallot	0	True	0	True
2.916260°	98.517670°	Shallot	0	True	0	True
2.939010°	98.508330°	Shallot	0	True	0	True
2.905250°	98.513330°	Shallot	0	True	0	True
2.898220°	98.500940°	Shallot	0	True	0	True
2.897640°	98.500770°	Shallot	0	True	0	True
2.896630°	98.499390°	Shallot	0	True	0	True
2.958620°	98.528440°	Shallot	-	False	-	False
2.926330°	98.524560°	Shallot	0	True	0	True
2.921510°	98.491420°	Tomato	1	True	1	True
2.974370°	98.538730°	Orange	1	True	1	True
2.974740°	98.539300°	Cabbage	1	True	1	True
2.963000°	98.535620°	Cayenne pepper	-	False	-	False
2.952030°	98.522790°	Cayenne pepper	1	True	1	True
2.952570°	98.522720°	Cabbage	1	True	1	True
2.937120°	98.528420°	Tomato	1	True	1	True
2.937290°	98.528560°	Tomato	1	True	1	True
2.895720°	98.526520°	Shallot	0	True	0	True
2.895680°	98.526540°	Shallot	0	True	0	True
2.920260°	98.524670°	Cayenne pepper	1	True	1	True
2.926480°	98.524580°	Tomato	1	True	1	True
2.926500°	98.524820°	Cabbage	1	True	1	True
2.921180°	98.507280°	Potato	1	True	1	True



a) Conditions of farm embankment on a scale of 1:50 m b) Classification of farm embankments on a scale of 1:50 m

Figure 5 Classification of GEE on farm embankments on a scale of 1:50 m.

combination of sentinel-1A and sentinel-2A images (Putri *et al.* 2018).

The accuracy of the location on the classification map was seen to weaken when zooming in with a scale of 1:50 m. The classification was not appropriate for placing the shallots, especially regarding area boundaries/garden embankments. This is because sentinel-2 is included in the medium-resolution satellite (Devara and Wijayanto 2021), so it lacks detail in small-scale images (Figure 5). This weakness can be overcome using high-resolution image data such as drone data (Sugara *et al.* 2022).

The Rapid Classification of Croplands method allows for faster area estimation than the eye estimation method. The eye estimate is an estimate of the harvest area using the eye view so that it has a great chance of subjectivity (Rusono 2019). The Rapid Classification of Croplands process begins with collecting coordinate data and plant classification. Calculating the land area is carried out with the help of computers/machine learning. This method can analyze the commodity balance before making policies related to shortages or oversupply (Sinaga and Hastuti 2019). A more accurate land area estimate will improve the integration of the shallot market (Zaeniyah 2022; Zain *et al.* 2022).

Estimating land area can also be done during outbreaks or disasters such as droughts and floods so that strategies can be carried out to mitigate climate anomalies (Limbong *et al.* 2020). The disadvantage of this method is that it requires sophisticated infrastructure, and data processing requires high specifications of computer devices, especially for a wider area (district or provincial level). In addition, in highland areas or during the dry season, many clouds will cover *the area of interest*. In principle, a clean image of clouds can be done by combining several images to improve areas covered by clouds (Papilaya 2022). The land area estimate requires a relatively short period to process the same conditions as the coordinate points from the field, so there are few options for clear images from clouds.

CONCLUSION

The Rapid Classification of Croplands method using *the* Google Earth Engine platform and image archive from Sentinel-2 can be an alternative method for determining the estimated area of shallot land. RF classification shows better accuracy than CART classification. The resulting accuracies were 72.13% (CART) and 71.33% (RF). The area accuracy was 95.2% (RF) and 24% (CART), and the location accuracies were 92% (CART and RF). Some of the limitations of this method are limited access and resolution, inability to describe the level of plantation management, and the

condition of the area covered by clouds, which will affect the accuracy of the results. The following research suggestion is to utilize drone technology in spatial data capture to reduce the limitations encountered in the use of sentinel-2 image archives.

REFERENCES

- Aspari R, Septianus TL, R, A. O. F. 2023. Rancang bangun mesin pengering skala miniatur dengan kapasitas 0,5 HP. *PISTON: Jurnal Teknologi*. 8(1): 1–9. <https://doi.org/10.55679/pistonjt.v8i1.14>
- Bofana J, Zhang M, Nabil M, Wu B, Tian F, Liu W, Zeng H, Zhang N, Nangombe SS, Cipriano SA, Phiri E, Mushore TD, Kaluba P, Mashonjowa E, Moyo C. 2020. Comparison of different cropland classification methods under diversified agroecological conditions in the Zambezi River Basin. *Remote Sensing*. 12(13). <https://doi.org/10.3390/rs12132096>
- BPS. 2023a. *Pengeluaran untuk Konsumsi Penduduk Indonesia per Provinsi 2023*. <https://www.archive.bps.go.id/publication/2023/10/20/b891f269877539ef8b0de1db/pengeluaran-untuk-konsumsi-penduduk-indonesia-per-provinsi--maret-2023.html>
- BPS. 2023b. *Statistik Pengeluaran Untuk Konsumsi Penduduk Provinsi Sumatera Utara 2022* (MP Samosir, CP Aridesyadi (eds.); BPS Sumatera Utara. <https://sumut.bps.go.id/publication/2023/07/12/fc5a9f6f5005a733734bb4cd/statistik-pengeluaran-untuk-konsumsi-penduduk-provinsi-sumatera-utara-2022-.html>
- BPS. 2023c. *Statistik Tanaman Hortikultura Provinsi Sumatera Utara 2022*. BPS Sumatera Utara. <https://sumut.bps.go.id/publication/2023/12/07/5f93b27f91659c9552bb1c0b/statistik-tanaman-hortikultura-provinsi-sumatera-utara-2022-.html>
- BPS Karo. 2023. *Kabupaten Karo dalam Angka (Karo Regency in Figures)*. <https://karokab.bps.go.id/publication/2023/02/28/1cb637693112bbe48cb6ab8b/kabupaten-karo-dalam-angka-2023.html>
- Devara T, Wijayanto AW. 2021. Machine learning applied to Sentinel-2 and Landsat-8 Multispectral and Medium-Resolution Satellite Imagery for the detection of rice production areas in Nganjuk, East Java, Indonesia. *International Journal of Remote Sensing and Earth Sciences (IJReSES)*. 18(1). 19. <https://doi.org/10.30536/j.ijreses.2021.v18.a3538>
- Firdaus M. 2021. Disparitas harga pangan strategis

- sebelum dan saat pandemi COVID-19. *Jurnal Ekonomi Indonesia*. 10(2): 107–120. <https://doi.org/10.52813/jei.v10i2.104>
- Flood N. 2013. Seasonal composite landsat TM/ETM+ Images using the medoid (a multi-dimensional median). *Remote Sensing*. 5(12): 6481–6500. <https://doi.org/10.3390/rs5126481>
- Limbong JJA, Ratu HH, Simbolon P, Prasetyo SYJ. 2020. Analisis indeks vegetasi area terdampak banjir bandang di Kabupaten Jayapura menggunakan metode clustering pada citra Landsat 8. *Indonesian Journal of Computing and Modeling*. 3(1): 1–6. <https://creativecommons.org/licenses/by/4.0/>
- Loukika KN, Keesara VR, Sridhar V. 2021. Analysis of land use and land cover using machine learning algorithms on google earth engine for Munneru river basin, India. *Sustainability*. 13(24). <https://doi.org/10.3390/su132413758>
- Marsuhandi AH, Soleh AM, Wijayanto H, Domiri DD. 2020. Pemanfaatan ensemble learning dan penginderaan jauh untuk pengklasifikasian jenis lahan padi. *Seminar Nasional Official Statistics*, 2019(1): 188–195. <https://doi.org/10.34123/semnasoffstat.v2019i1.247>
- Mulyaqin T, Kardiyono K, Hidayah I, Ramadhani F, Yusron M. 2022. Deteksi alih fungsi lahan padi sawah menggunakan Sentinel-2 dan Google Earth Engine di kota Serang, Provinsi Banten. *Jurnal Ilmu Pertanian Indonesia*. 27(2): 226–236. <https://doi.org/10.18343/jipi.27.2.226>
- Musfiza D, Armi I, Arini D, Fikri S. 2023. Aplikasi penginderaan jauh untuk pemetaan lahan sawah. *Jurnal Teknik Indonesia*. 2(April): 14–28. <https://doi.org/10.58860/jti.v2i2.19>
- Noi Phan T, Kuch V, Lehnert LW. 2020. Land cover classification using google earth engine and random forest classifier-the role of image composition. *Remote Sensing*. 12(15). <https://doi.org/10.3390/RS12152411>
- Papilaya PPE. 2022. Aplikasi Google Earth Engine dalam menyediakan citra satelit sumber daya alam bebas awan. *Makila*. 16(2): 96–103. <https://doi.org/10.30598/makila.v16i2.6586>
- Permatasari NKF, Tambunan MP, Mannesa MDM, Tambunan RP. 2021. Pengaruh kekeringan pada produksi tanaman padi di Kabupaten Majalengka dengan penginderaan jauh Metode Ndvi. *Jurnal Geosaintek*. 7(1): 17. <https://doi.org/10.12962/j25023659.v7i1.8205>
- Putri DR, Sudarsono B, Awaluddin M. 2018. Analisis kombinasi citra Sentinel-1a dan citra Sentinel-2a untuk klasifikasi tutupan lahan (Studi kasus: Kabupaten Demak, Jawa Tengah). *Jurnal Geodesi Undip*. 7(2): 85–96.
- Rahayu S, Wanita PW, Kobarsih M. 2014. Penyimpanan benih padi menggunakan berbagai jenis pengemas. *Agrin*. 15(1): 36–44.
- Rusono N. 2019. Improvement of Rice food data for quality food development planning. *The Indonesian Journal of Development Planning*, 3(2): 253–265. <https://doi.org/10.36574/jpp.v3i2.78>
- Samad MY. 2016. Pengaruh penanganan pasca panen terhadap mutu komoditas hortikultura. *Jurnal Sains Dan Teknologi Indonesia*. 8(1): 31–36.
- Septiana LR, Machfud, Yuliasih I. 2017. Peningkatan Kinerja rantai pasok bawang merah (Studi kasus: Kabupaten Brebes). *Jurnal Teknologi Industri Pertanian*. 27(2): 125–140. <https://doi.org/10.24961/j.tek.ind.pert.2017.27.2.125>
- Serra T, Gil JM. 2013. Price volatility in food markets: can stock building mitigate price fluctuations? *European Review of Agricultural Economics*. 40(3): 507–528. <https://doi.org/10.1093/erae/jbs041>
- Sinaga BM, Hastuti AF. 2019. Dampak kebijakan impor dan faktor eksternal terhadap kesejahteraan produsen dan konsumen bawang merah di Indonesia. *Journal of Agriculture, Resource and Environmental Economics*. 2(1): 38–53. <https://doi.org/10.29244/jaree.v2i1.25963>
- Sugara A, Siregar VP, Agus SB. 2022. Klasifikasi habitat benthik perairan dangkal dari citra Worldview-2 menggunakan data *in-situ* dan drone. *Jurnal Ilmu dan Teknologi Kelautan Tropis*. 19(2): 135–150. <https://doi.org/10.29244/jitkt.v12i1.26448>
- Tamiminia H, Salehi B, Mahdianpari M, Quackenbush L, Adeli S, Brisco B. 2020. Google Earth Engine for geobig data applications: A meta-analysis and systematic review. *ISPRS Journal of Photogrammetry and Remote Sensing*. 164(April): 152–170. <https://doi.org/10.1016/j.isprsjprs.2020.04.001>
- Zaeniyah F. 2022. Integrasi vertikal pasar bawang merah di Provinsi Jawa Timur. *Jurnal Ekonomi Pertanian dan Agribisnis* (JEPA). 6(4): 1583–1596. <https://doi.org/10.21776/ub.jepa.2022.006.02.2>
- Zain AI, Widjojoko T, Mamdamdari AN. 2022. Variasi harga dan integrasi pasar vertikal bawang merah di Kabupaten Brebes. *Jurnal Ekonomi Pertanian dan*

Agribisnis. 6(2): 348–359.
<https://doi.org/10.21776/ub.jepa.2022.006.02.2>

Zulkarnain M, Marsisno W. 2022. Penerapan pembelajaran mesin untuk estimasi luas lahan bawang merah berdasarkan data citra satelit resolusi

menengah: Studi kasus di Kabupaten Nganjuk, Jawa Timur (Application of machine learning for estimating the area of shallots planted based on medium. *Seminar Nasional Official Statistics 2022*. 1005–1016. <https://doi.org/10.34123/semnasoffstat.v2022i1.1307>