Estimating Oil Palm Production Using Sentinel 2: By-age Separation Method

Muhammad Adnan Shafry Untoro[1*], Lina Indriani[1], Bagus Wiratmoko[1], Ahnaf Hampar Sasongko[1], Balada Wisma Sandika[1], Donna Rizky Amora[1], Nur Lely Wahyunissa[1] and Sigit Heru Murti B S[1]

[1]Department of Geographic Information Science, Faculty of Geography, Gadjah Mada University, Yogyakarta 55281, Indonesia

muhammad.adnan.s@mail.ugm.ac.id

Abstract. Multispectral imagery can be used for estimating production of oil palm fruit in a plantation. One of the multispectral imagery that capable for production estimation is Sentinel 2. Sentinel 2 is a free imagery that have great 10 m medium spatial resolution bands, which are red, green, blue, and near-infrared. These bands can be used to generate several generic vegetation index such as Ratio Vegetation Index (RVI) and normalized difference vegetation index (NDVI) also non-generic vegetation index such as modified soil adjusted vegetation index (MSAVI) and atmospherically resistant vegetation index (ARVI). Those vegetation index applied for estimating the production of oil palm fruit in PT. Perkebunan Nusantara VII's plantation at Rejosari, Natar. By-age separation method demonstrate that it has a better accuracy which is 87.5 % rather than normal method that fully transformed without being separated by-age. The fully transformed method accuracy are 61.1 % for RVI, 59.1 % for NDVI, 55.3 % for MSAVI, and 62.1 % for ARVI.

Keywords: Estimation of crop production, Oil palm, Vegetation Index, Sentinel 2

1. Introduction

Oil palm or *Elaeis guineensis jacq* is one of the most widely produced and consumed commodities around the world. Oil palm is also the most profitable oil crops commodities because it's gives highest ratio of production per hectare [4]. Indonesia is the biggest palm oil producer and exporter in the world that in 2015 the total area of oil palm plantation in Indonesia reaches 11.3 million hectares and producing 31.28 million tons of palm oil [1]. East coast of Sumatera and Aceh is the area of first oil palm plantation taking place [6] and then it spread out to South Sumatera, West Sumatera, Riau, Bengkulu, Jambi, Lampung, West Java, West Kalimantan, Central Kalimantan, East Kalimantan, Maluku, Sulawesi, and Papua [8]. PT Perkebunan Nusantara is one of company that carry out oil palm plantation of Indonesian government in 3 provinces.

Currently, the calculation of the estimated amount of palm oil production is mostly done terrestrially. The terrestrial method certainly takes a lot of time and energy because the oil palm plantations usually are so large that this method is considered inefficient. Remote sensing can be used to assist the calculation of production estimation. The estimation of palm oil production in remote sensing requires

data of tree's age, vegetation index, and heavy sample data of palm oil in the field [7]. The data is processed together with the vegetation index transformation of the remote sensing imagery used, which is Sentinel 2. Therefore, calculation of production estimation is no longer done by census in the field (terrestrial) but uses sampling technique and then integrated with remote sensing imagery. However, estimating production of oil palm using remote sensing can be challenging because oil palm tree physical appearance is almost identical at age of 5 – 25, but they have different level of production. Therefore, a method which can differentiate oil palm tree physical appearance and relate it with their level of production is needed.

2. Study Site

The research location of palm oil production estimate is done at oil palm plantation owned by PT. Perkebunan Nasional VII Lampung Rejosari Unit. The plantation is divided into four smaller part called *afdeling,* a Dutch word, which can be translated into section. Each section divided into smaller part called block. The plantation is located on Rejosari, Lampung Selatan. PT. Perkebunan Nasional VII oil palm plantation has variety in age. The oldest trees planted on 1992 and the youngest was planted on 2010.

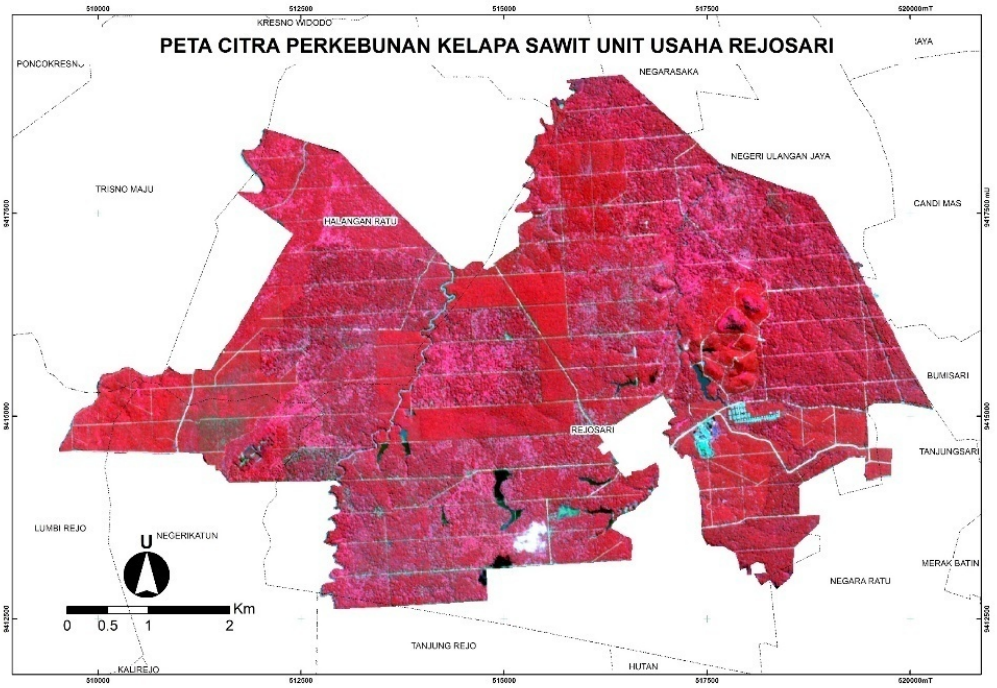

Figure 1. Overall look of PT. Perkebunan Nusantara VII Oil Palm Plantation from Sentinel 2 imagery in false color composite. Vegetation is red and building is blue.

3. Material and Methods

3.1. Data

3.1.1. Sentinel 2. The main data used is Sentinel 2 dated June 22th, 2016. Dark object subtraction using Semi-Automatic Classification Plugins (SCP) on Quantum GIS is being done to minimize the noise of atmospheric effect [2]. SCP correction is based of ESA's Sentinel 2 formula [3]. SCP basically is a shortcut to help people doing atmospheric correction for the imagery.

The imagery then masked into only study area site. Corrected and masked image then transformed with vegetation index. The goal of this transformation is to highlight the information of vegetation objects on the surface of the earth. The vegetation index used in this paper divided into two group, which are generic and non-generic index. The generic index used are generic vegetation index such as *Ratio*

Vegetation Index (RVI) and *normalized difference vegetation index* (NDVI). The non-generic vegetation index used are *modified soil adjusted vegetation index* (MSAVI) to reduce soil background and *atmospherically resistant vegetation index* (ARVI) to reduce atmospheric effect [5].

3.1.2. Plantation Time Map. Used to mask Sentinel 2 Imagery and separate the oil palm blocks based on different planting ages. The reason for the separation is because the age difference of oil palm is have different production capability [4]. The image transformation of vegetation index is explained into 3 classes of planting year i.e. 1992-1998, 1999-2005, and 2006-2010.

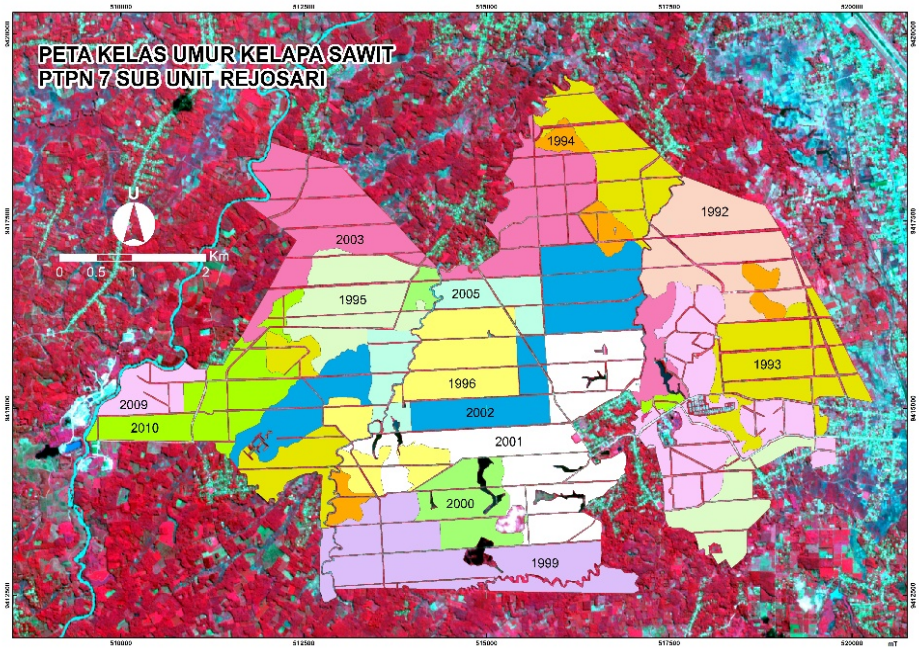

Figure 2. PT. Perkebunan Nusantara VII Plantation Time Map, oldest trees planted on 1992 and the youngest was planted on 2010.

3.2. Methods

3.2.1. By-age separation method. This age separation method is inspired from the work of Setyowati (2015) that build oil palm production model based on a multivariate regression analysis from age of the tree, vegetation index value, and the production itself. In this paper, the age of the tree separated from the regression analysis so the modelling is need to be done age-by-age.

3.2.2. Regression analysis. Regression analysis used is linear regression analysis which is done on each age classification, so that yielded some regression equation and some image of model produced. The simple linear regression algorithm used is as follows [10]:

$$Y = a + bX \quad (1)$$

Y = Production value according to empirical model (dependence variable)
a and b = Coefficients of values generated at model
X = Image pixel value used to build the models (independence variable)

3.2.3. Sample determination and field sampling. Sample determination was done by using stratified random sampling method. The data taken in the field is the data weight of fruit in one rotation in one unit of analysis in the form of 1.5 times bigger than the pixel of Sentinel 2 to avoid GPS location error. The number of samples that were planned before the field were 30 samples but due to time constraints then the sample can be measured when in the field as many as 18 samples.

3.2.4. Accuracy test. The accuracy test of production estimation result is based on the difference between the total productions of the modelling results from Sentinel 2 image compared to the production data of PT. Perkebunan Nusantara VII Rejosari Business Unit in six months where the production is start to declining because when the samples taken is when the production declined. Six months is time required for oil palm to be formed [4]. February-July 2016 is the time when PT. Perkebunan Nusantara VII oil palm production is declining. The accuracy test method refers to Susetyo (2010) production estimation and production of reality which can then be calculated the level of accuracy. Estimation test formula estimation model used is as follows:

$$Accuracy = 100\% - ((\frac{estimation-reality}{reality})100\%) \quad (2)$$

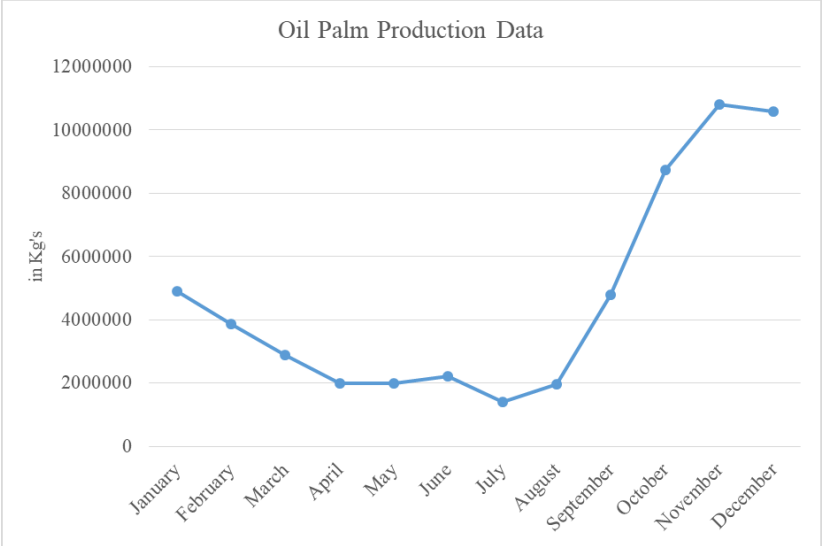

Figure 3. Production Data of Oil Palm in PT. Perkebunan Nusantara VII.

4. Results

4.1. Normality test for data samples

The samples is needed to build the production model. Model development is done to correlate between indices generated from image with existing production in the field. Komogolrov-Smirnof method is used to testthe normality of samples data by determining the value of S(n), Z-factor, F(x) dsn difference on a sample training data set. Normality test is done by comparing the maximum difference value (Dn = max) with the maximum difference resul from KomogolrovSmirnof test (Dn, a). From 18 samples acquired a Difference Maximum value is 0.61396, while the difference value is of 0.307 so it can be concluded that the overall data is normally distributed.

Table 1. Production Data of Oil Palm in PT. Perkebunan Nusantara VII

Sample number	X (m)	Y (m)	Afdeling	Block	Plantation time	Age classification
1	519379	9416486	4	3	1992	1 (1992 -1998)
2	519335	9416475	4	3	1992	
3	512482	9414570	2	4	1993	
4	514470	9415941	3	2	1996	
5	514209	9416012	3	3	1996	
6	515376	9414751	1	2	2001	2 (1999-2005)
7	515379	9414770	1	1	2002	
8	515525	9415781	3	1	2002	

Sample number	X (m)	Y (m)	Afdeling	Block	Plantation time	Age classification
9	515531	9415764	3	1	2002	
10	512483	9414640	2	3	2002	
11	515097	9416836	3	4	2005	
12	517665	9414350	4	1	2009	
13	517722	9414181	4	1	2009	
14	519247	9414666	4	2	2009	
15	519279	9414723	4	2	2009	3 (2006-2010)
16	511851	9414945	2	1	2010	
17	511805	9414936	2	2	2010	
18	512260	9415363	2	2	2010	

4.2. Oil Palm Production Model

Production model is built based on the regression formula. Each model is compared with PT. Perkebunan Nusantara's production data to justify the best index and method. There are two methods of regression. The first method is a method that does not pay attention to the planting time so that regression is done using all sample data. Estimated oil palm production on all vegetation indices regardless of planting age group overestimate. Secondary data shows that the production of palm oil plantation in February – July 2016 was 14,300.9 tons; while the estimation result using RVI vegetation index was 19,860.5 tons; NDVI 20,139.7 ton; MSAVI 20,691 tons; and ARVI 19,715,3 tons. The second method, by-age separation, implies that oil palm have different reaction to each vegetation index based on their age, so the best way to predict oil pam production is to use different vegetation index for each age class. This method done by separating the plantation site into four age classes; 1992-1996, 2001-2005 and 2009-2010 then each age classes will be modelled using each indices regression independently. The most accurate model from each age classes then merged into one imagery model so called by-age separation model. This method result is 16,228.9 tons of oil palm fruits from February to July 2016.

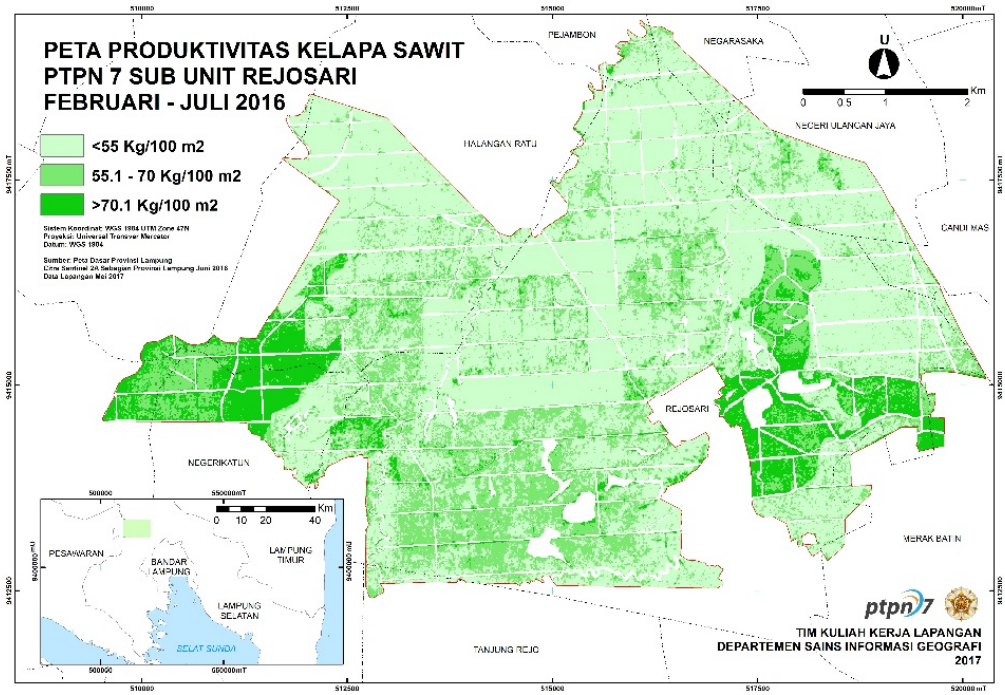

Figure 4. By-age separation method production model. This model is built from best estimation model in each age classes and predicted 16,228.9 tons of oil palm fruits from February to July 2016.

5. Discussion

The number of samples taken on the field to build the production models is actually 22 samples, while used in the model builder to determine the production is only 18 samples. There are 4 samples removed because it was too different from the existing training data.

Based on the treatment with the by-age separation method, the RVI index has the highest R square value in the first group (1992-1996) and the second (1999-2005) with the values of 0.4592 and 0.4787 while in the third group (2009-2010) The best index is ARVI with a value of 0.0344. The second method, the method that does not pay attention to age group, the highest R square value is at ARVI index with value 0.2495. Regression conducted using field data and image data, where imagery data will affect the level of productivity of oil palm in the field. The graph below shows the best regression results between the samples and the vegetation index image in each age group.

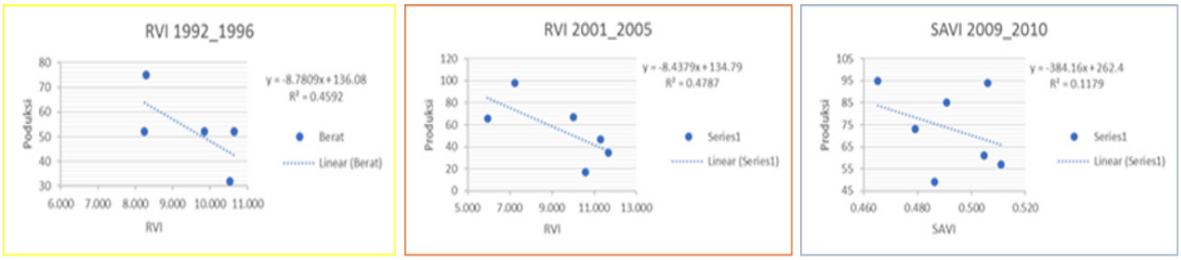

Figure 5. Best vegetation index in each age group to build by-age separation model

The best transformation index for the 1992-1996 planting age group was ARVI with 65.9% accuracy and estimated production of 5,961.8 tons. The best transformation index for the 1999-2005 planting year was RVI with an accuracy of 97.9% and an estimated production of 8,059.0 tons. The best transformation index for the 2009-2010 planting group is NDVI with an accuracy of 63.81% and an estimated production of 2,208.1 tons.

The accuracy test of production estimation result is done based on the difference between the total production of the modelling result from Sentinel 2 imagery compared with the production data of PTPN 7 Rejosari Business Unit in February-July 2016. The result of accuracy test shows that the best accuracy is on the modelling with separation of planting age group with level Accuracy of 87.5%.

6. Conclusion

By-age separation method is better than the fully transformed without being separated by-age. It is because which can differentiate oil palm tree physical appearance and relate it with their level of production. By-age separation method implies that oil palm have different reaction to each vegetation index based on their age. The best transformation index in the estimation of oil palm production planting year 1992-1996 is ARVI, in 1999-2005 planting year that is RVI, and in planting year 2006-2010 that is NDVI. The combined production estimates for each age group resulted in a total production estimate of 16,228.9 ton/semester with an 87.5% modeling accuracy.

References

[1] Badan Pusat Statistik 2015 *Statistik Kelapa Sawit Indonesia* (Jakarta: Badan Pusat Statistik)
[2] Chavez P S 1996 *Image-based Atmospheric correction-revisited and improved (Photogrammetric Engineering & Remote Sensing vol 62)* p 1025-36.
[3] Congedo Luca 2017 *Semi-automatic classification plugin documentation*
[4] Corley R H V and P B Tinker 2003 *The oil palm 4th edition* (Victoria: Blakwell)
[5] Danoedoro P 2012 *Pengantar penginderaan jauh digital* (Yogyakarta: ANDI Yogyakarta)
[6] Fauzi Yan Y.E. Widyastuti I. Setyawibawa and R. Hartono 2002 *Kelapa Sawit: Budidaya, Pemanfaatan Hasil Limbah, Analisis Usaha, dan Pemasaran (Revisi)* (Jakarta: Penebar Swadaya)
[7] Setyowati H A 2015 *Aplikasi citra SPOT-6 berbasis transformasi indeks vegetasi untuk estimasi produksi kelapa sawit (Elaeis guneensis jacq.) (kasus perkebunan kelapa sawit PT. Tunggal*

Perkasa Plantations, Air Molek, Kabupaten Indragiri, Hulu, Provinsi Riau, Sumatera.) Essay (Yogyakarta: Fakultas Geografi UGM)
[8] Sunarko 2007 *Petunjuk praktis budidaya dan pengolahan kelapa sawit* (Jakarta: Agromedia Pustaka)
[9] Susetyo Imam 2012 *Pemodelan estimasi produksi tanaman karet (Havea brasiliensis mell.-Arg) berbasis citra satelit, potensi genetik, dan data satuan medan (studi kebun Getas, PTPN IX Salatiga, Kabupaten Semarang, Jawa Tengah)* Thesis (Yogyakarta: Fakultas Geografi UGM)
[10] Pranowo 1982 *Statistik praktis* (Yogyakarta: Ananda Yogyakarta)

Acknowledgements
This research was part of Department of Geographic Information Science's 3rd field work and supported by PT. Perkebunan Nusantara VII who allow us to done our field work in their plantation and provided secondary data. We also like to thank Faculty of Geography, Universitas Gadjah mada who help us to fund this publication.