THE EFFECT OF MINIMUM NOISE FRACTION ON MULTISPECTRAL IMAGERY DATA FOR VEGETATION CANOPY DENSITY MODELLING

A. M. Syarif[a], I. S. W. Kumara[a]

[a] *Department of Geographic Information Science, Faculty of Geography, Universitas Gadjah Mada 55281, Indonesia*

Abstract: *Minimum Noise Fraction (MNF) is known as one of the methods to minimize noise on hyper spectral imagery. In addition, there are not many studies have tried to show the effect of MNF transform on multispectral data. This study purposes to determine the effect of MNF transform on the accuracy level of vegetation density modeling using 10 meters Sentinel-2A spatial resolution (multispectral data) and to know the cause. The study area is located in parts of Sapporo City, Hokkaido, Japan. Vegetation density is modeled through vegetation index approach, Normalized Difference Vegetation Index (NDVI). The results show that the coefficient correlation of vegetation density data and vegetation index regression after MNF transformation (0.801623) has higher value than the same regression without the MNF (0.794481). However, better correlation does not represent the better accuracy on vegetation density modeling. Accuracy calculation through standard error of estimate shows the use of MNF in multispectral data for vegetation density modeling causes the decrease of model accuracy value. The accuracy of vegetation density model without involving MNF transformation reached 91.4%, while the model accuracy through MNF transformation before vegetation density modeling reached 90.89%. The insignificant increased accuracy is occurred due to the limited number of multispectral image information compared to hyper spectral image data.*

1. INTRODUCTION

Remote sensing data had been widely used to identify object or phenomenon in earth surface, includes environmental problem, environmental management, and regional planning. Technological developments in images data still leave problems related to data quality that decreases due to noise. The presence of noise causes a loss of the real image spatial resolution (Letexier & Bourennane, 2008), furthermore can reduce the quality of object detection through image classification and segmentation (Ebadi et al., 2013). Noise interferences that decreasing the remote sensing data performance for transformation and analysis are necessarily to be reduced. For decades, many algorithms and methods have been developed to remove the noise, one of those is through Minimum Noise Fraction (MNF). MNF commonly used to reduce noise in remote sensing data, especially in hyper spectral images. High dimension data in hyper spectral image may reduce the efficiency of data processing, aside from the information mining process. The more spectral bands possessed by an image data, more training samples needed to obtain an acceptable classification accuracy, therefore begin unfavorable in image processing. Dimensionality reduction is one of effective solution to overcome the problem (Jia et al., 2013). MNF is able to create a set of good-quality image data from noise-contained data. The result producing more reliable noise identification and removal, also allows the preservation of components that are considered to contain useful information (Luo et al., 2016). MNF is a linear transformation method consists of two separate PCA rotations and noise removal steps (Green et al., 1988).

Although noise reduction had been reported as an effective method on hyper spectral data, the efficiency on the multispectral data have not been tested explicitly. Because of the limited knowledge about the efficiency, some question arises when MNF is applied to multispectral image. This research attempts to analyze the effect that occurs when MNF is applied to multispectral data. On the contrary of hyper spectral data dimensionality, multispectral image has limited bands and information that would be

also reduced through the method to some extent, the MNF effectiveness would be declined greater. The research of MNF through hyper spectral airborne image data to classify land cover has been conducted (Frassy et al., 2013), however the effect of MNF analysis is done by comparing the accuracy of spatial modeling results of vegetation canopy density. The vegetation canopy mapping topics is interesting and challenging considered to the lack of research related to the influence of MNF usage analysis of vegetation canopy density model (quantitative stratified data) which is not included as categorical data. The simple analysis of MNF effect carried out of by this study is expected to be able to describe influence of MNF to multispectral data.

2. STUDY AREA AND IMAGE DATA

2.1. Study Area

This study conducted in Sapporo, capital city of Hokkaido Prefecture, Japan. Sapporo city included to mid-latitude region of with cold climate at the northern tip of Japanese islands (Figure 1a). Although Sapporo city is including as crowded city, the environment quality is still considered good. Well-maintained environment area and the number of open green space indicate of citizen awareness of environmental impact of people's health and life. Figure 1b show the scattered open green space scattered throughout Sapporo City and surrounding areas, whereas decent to be used to identify the influence of MNF towards the accuracy of vegetation canopy density modeling.

2.2. Image Data

Sentinel-2A is a part of Sentinel mission by European Space Agency (ESA), available freely since June 2015 at http://scihub.copernicus.eu/dhus/. Complete Sentinel-2A dataset has 13 bands; 4 bands of 10-meter resolution; 6 bands of 20-meter resolution; and 3 bands of 60-meter resolution. Although Sentinel-2A data considered as recently released, many geospatial applications have been carried out by this images data such as urban area assessment (Pesaresi et al., 2016), forest monitoring (Navarro et al., 2017), minerals detection (Chen et al., 2017), and coastal and marine resources (Fauzan et al., 2017; Traganos & Reinartz, 2017). As many potentials of usage possibility through Sentinel-2A remote sensing approach, the vegetation canopy density modeling in this study may be done. Figure 1b below is natural composite bands of Sentinel-2A recorded at May 23, 2016 in the Sapporo city and surrounding. We only utilize 4 of visible bands at 10-meter resolution (blue, green, red, and near infrared wavelength) that sufficient enough to detect vegetation canopy.
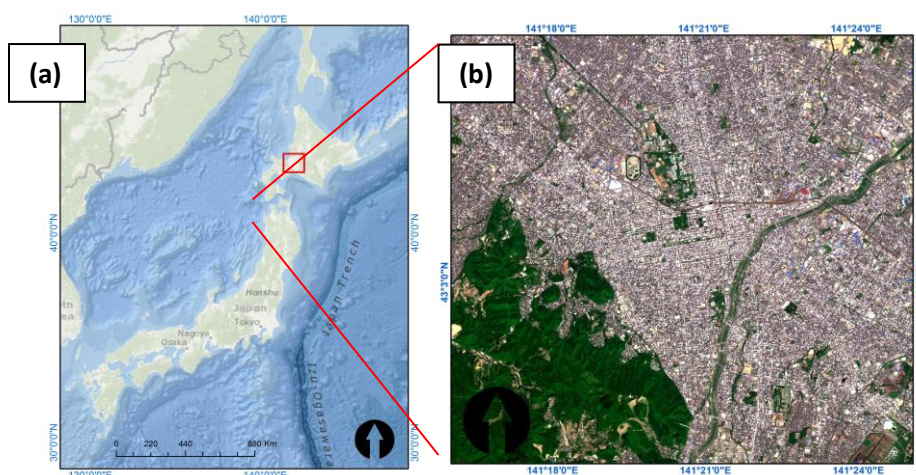

Figure 1. Study area (a) Japanese Islands and (b) Sapporo City and surrounding by Sentinel-2A image

3. METHODS

3.1. Data Pre-processing

Vegetation canopy density modeling requires systematic image data pre-processing through radiometric to atmospheric correction. Radiometric correction is used to correct the image pixel value to fit with object surface reflectance value, while atmospheric correction is used to minimize the atmospheric disturbance of reflected energy from the object surface to sensor. Fast Line-of-sight Atmospheric Analysis of Spectral Hyper cubes (FLAASH) module used to correct Sentinel-2A data to suppress the effects of water

vapor, oxygen, carbon dioxide, methane, ozone, and aerosols (Danoedoro, 2012). Actually, ESA have been developed Sentinel-2A reflectance value transformer up to bottom-of-atmosphere (BOA) image data using Sen2Cor plugin on SNAP application, but we trying to correcting the image data using FLAASH module.

First step for Sentinel-2A radiometric correction is convert pixels from the TOA reflectance value (L_{TOA}) into TOA radiance value (R_{TOA}) using formula (1) below.

$$R_{TOA} = [\: U * L_{TOA} * ESUN * (\cos \theta_Z) \:] / \pi \qquad (1)$$

where U is earth-sun distance corrections factor ($1/d^2$), L_{TOA} is TOA reflectance value, ESUN is solar irradiance value of each band, θ_Z is sun zenith angle value, and π is 3,14159. After transform the image pixel into radiance value, the FLAASH module is operated to suppress the atmosphere effect on pixels. FLAASH require BIP or BIL type data as input. The visibility value of the surrounding area and some parameters, i.e. sensor altitude, ground elevation, flight time, and aerosol height are need to be provided for the algorithm process. Technical procedure of FLAASH method for Sentinel-2A data retrieved from Lantzanakis et al. (2016).

3.2. Minimum Noise Fraction (MNF)

MNF transform is used to determine the inherent dimensionality of image data, to segregate noise in the data, and to reduce the computational requirements for subsequent processing (Boardman & Kruse, 1994). MNF is executed by using the ENVI software, begins with Forward MNF. The forward MNF of 4 Sentinel-2A multispectral bands is set to produce 4 MNF images. The last forward MNF output band is considered containing a lot of noise and removed at inverse MNF process. Band 1, 2, and 3 of forward MNF are returned initial spectral bands dataset (4 bands) using inverse MNF to produce noise-reduced Sentinel-2A image data. Figure 2 show detailed procedure of MNF transformation.
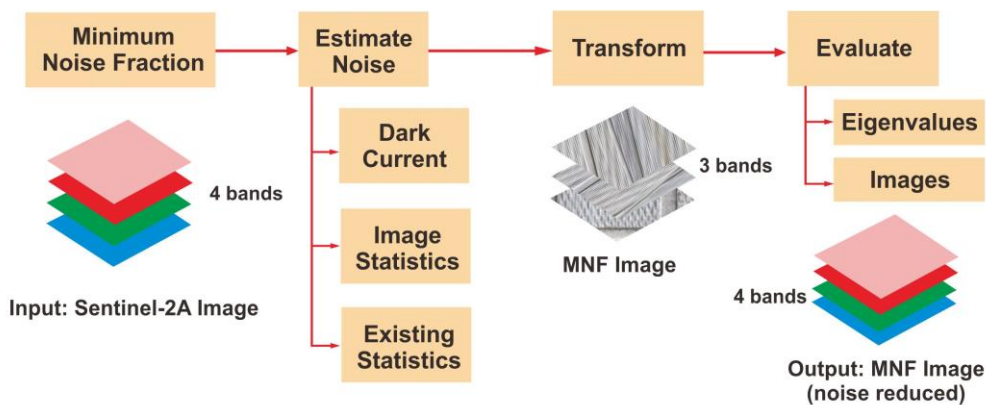

Figure 2. MNF Procedure (Exelis, 2015) with modifications

3.3. Vegetation Index Transformation

NDVI image transformation method provides optimum results to detecting the vegetation variation in relation to its density (Bhandari et al., 2012). NDVI principle based on the comparison between the measured vegetation canopy reflections on the red band and near infrared band. The difference in NDVI value intervals is used to simplify the classes of recorded objects. The NDVI formula (2) is derived from Rouse et al. (1974) as follows.

$$NDVI = \frac{B_{NIR} - B_{Red}}{B_{NIR} + B_{Red}} \qquad (2)$$

where B_{NIR} is near infrared band in Sentinel-2A data (Band 8) and B_{Red} is red band in Sentinel-2A data (Band 4)'

3.4. Vegetation Canopy Density Modeling

The vegetation canopy density modeling was derived from the relation between the vegetation index value parameter on the image pixel and the obtained field vegetation canopy density sample data. The field vegetation canopy density estimated using card on Figure 3. Linear regression adapted to linking the modeled NDVI data and field data. The regression equation value of the NDVI as dependent variable and the field canopy density as independent variable is extrapolated the pixel value throughout the whole study area into estimated canopy density. High correlation value of both data indicates the developed model is well enough to estimate the vegetation canopy density. Correlation goodness described from the coefficient of determination value of the regression results.
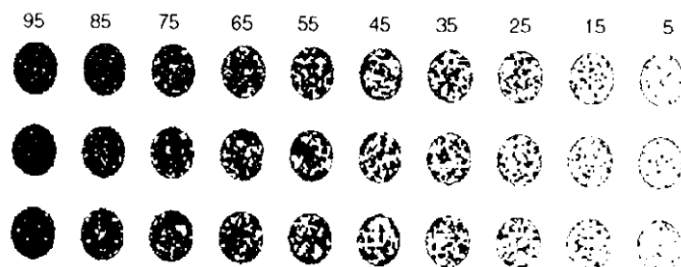
Figure 3. Field Vegetation Canopy Density Estimation Card

4. RESULTS AND DISCUSSION

4.1. MNF Transformation Result

MNF was applied on 4 Sentinel-2A bands (blue, green, red, and near infrared) after FLAASH correction was performed. The MNF-transformed image produces 4-band data that have variation in information and noise contents each other. First band of MNF image contain higher information than band 2, 3, and 4. On the contrary, fort MNF band had least information content and highest noise than another band. In this case, the MNF image not divides the image based on spectral band, but based on the content of information in the whole image.

Forward MNF generate 4 bands with specific information content levels (Table 1). MNF band 1 contains 71.52% of information, MNF band 2 contains 15.65% of information, the MNF band 3 contains 8.75% of information, and the MNF band 4 contains 4.08% of the information on the whole image information. The processing results indicate that MNF band 4 has the most noise. The abundant of noise and insignificant information on the MNF band 4 made the band unusable in subsequent process. Thus, the generate Sentinel-2A image after forward MNF has minimized noise content (information up to 95.92%).

Table 1. Information Content Level on MNF image

MNF band	Percent Information	Information Acc.
1	71.52 %	71.52 %
2	15.65 %	87.17 %
3	8.75 %	95.92 %
4	4.08 %	100 %

Unused MNF band 4 was not involved in the MNF Inverse process, so only applied to the three higher-information bands. The MNF inverse process performed to return image sequences based on spectral bands without noise, not based on information content levels like forward MNF bands. Inverse MNF converts 3 MNF bands into 4 image bands (blue, green, red, and near infrared). The image quality has been improved because of the eliminated noises, hence particular object feature processing and analysis bringing more accurate results (Frassy et al., 2013).

4.2. Normalized Difference Vegetation Index (NDVI) Result

The NDVI transformation is applied to image that has undergone the MNF and without MNF. This transformation is used to assist quantities vegetation data extraction in an image pixel. The NDVI transformation equation involving two bands of Sentinel-2A, red and near infrared. The NDVI

transformation results show the different range of values in image without MNF and image that has undergone the MNF process. Figure 4a show the NDVI image values without MNF range from -0.258 to 0.887. On the other side, Figure 4b show the NDVI image values with MNF range from -0.247 to 0.879.
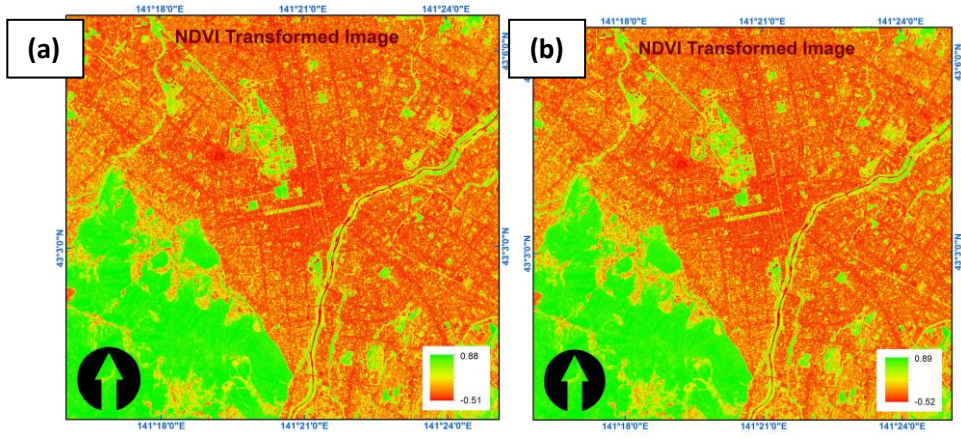

Figure 4. NDVI Map (a) without MNF and (b) with MNF

NDVI transformation results with and without MNF (Figure 4) were insignificantly different relying on its visual properties, also due to the insignificant NVDI value difference. In general, high NDVI values indicate the presence of large amounts of vegetation in a pixel. Visually, the vegetation presence shows on green color pixel, bare soil and build-up area on red color, and water body on yellow color. The vegetation's are scattered throughout the study area, majority found in southwestern area of hills and forest also dominated with woody vegetation (Bhandari et al., 2012). Another minor vegetation area on the river banks that pass through the city and some open green space within the city (especially the Hokkaido University campus area and surroundings).

4.3. Vegetation Canopy Density (VCD) Modeling

VCD estimation is conducted using empirical model of NDVI value and field canopy density. The field canopy density data divided into two, the model sample points and the accuracy assessment sample points that obtained using the density estimation card (Figure 3). There are 30 points of model sample and 8 points of accuracy sample distributed within the study area (Figure 5).
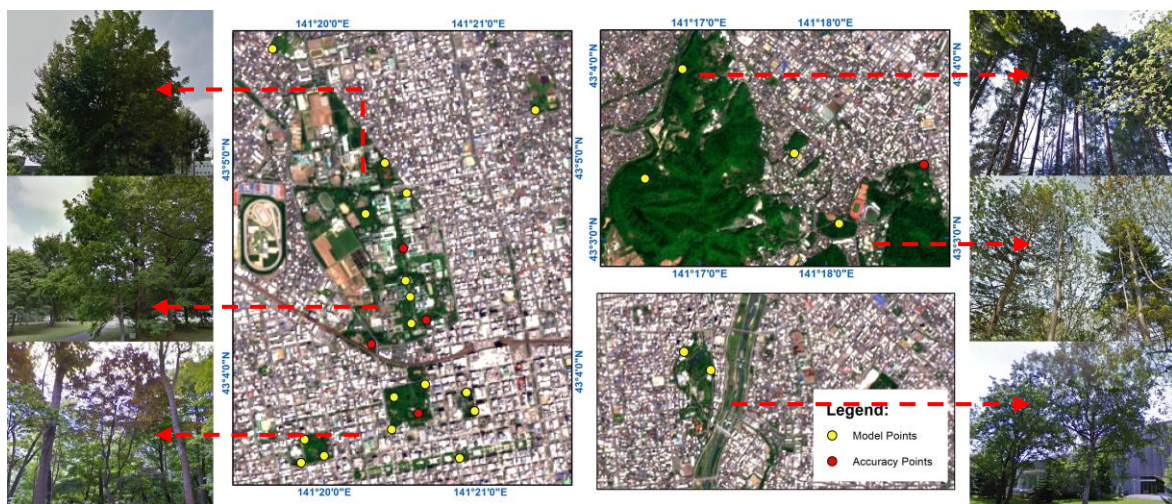

Figure 5. Model and Accuracy Sample Distribution Map

Regression between NDVI and field vegetation canopy density resulting the emerge of coefficient of determination 0.794481 (R^2 = 0.6312) for image without MNF (Figure 6a) and 0.801623 (R^2 = 0.6426) for image with MNF (Figure 6b). The R^2 value for MNF transformation is slightly higher than the other image

data. The two variables relations of NDVI and canopy density sample are highly categorized, able to represent the vegetation canopy density distribution (Bhandari et al., 2012). Through regression, the equations for the pixel value inversion to vegetation canopy density modeling were produced. The vegetation canopy density model was carried out by inputting the equations into statistical analysis.
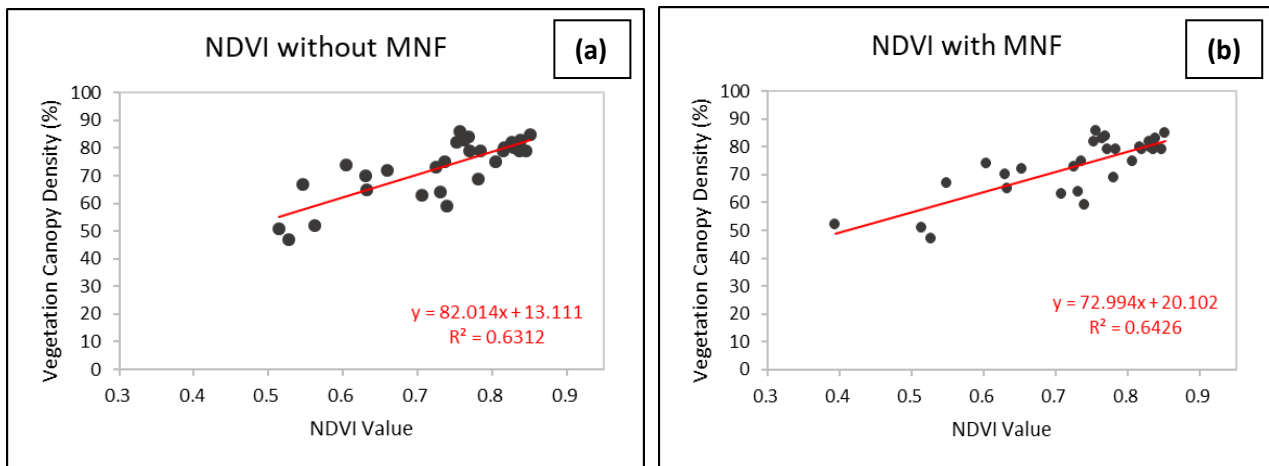
Figure 6. Regression between NDVI and Vegetation Canopy Density (a) without MNF; (b) with MNF

The two-derived vegetation canopy density models were measured to gain the model accuracy each. Accuracy was carried out through standard error calculation on 8 points of sample accuracy. The result of accuracy assessment show that vegetation canopy density model using MNF transformation is 90.889%, meanwhile the model without using MNF is 91.402%. The accuracy using MNF transformation is slightly lower than the original image transformation. Noise reduction in multispectral images does not fully improve the quality of canopy density modeling results, rather below the unremoved noise image. This condition was due to the small number of Sentinel-2A multispectral image bands (4 bands). The vegetation canopy density models derived from the regression transformation has 1.0% − 2.0% density difference. Even for the visual of both vegetation densities map not slightly different (Figure 7).
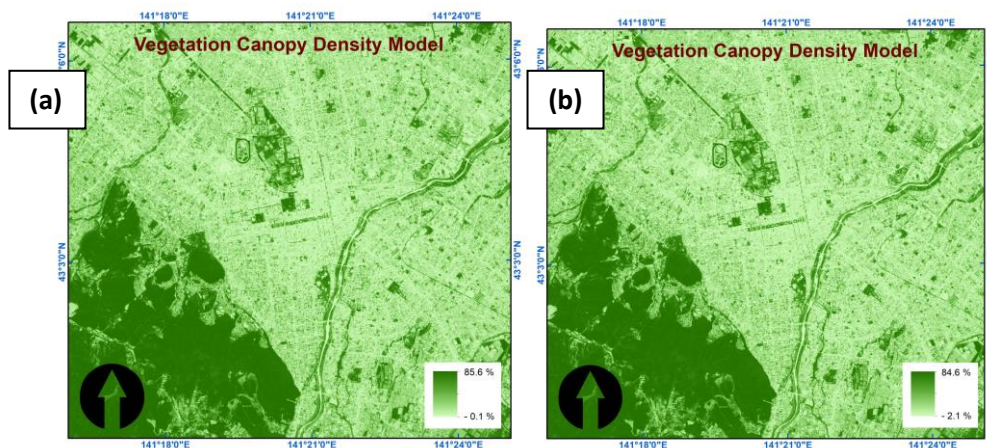
Figure 7. Vegetation Canopy Density Map (a) without MNF and (b) with MNF

MNF procedure not only removes noise in image data, but also the part of information due to PCA mechanism inside MNF algorithm (Luo et al., 2016). Band elimination for noise removal purpose has significant impact of reducing the information within the image. Mention to Table 1 that show the availability of information on each forward MNF band. The MNF images from Sentinel-2A only consist of 4 bands, cause the last band still preserve a great amount of information (4.08%). According to the MNF processing result, the declination of information contents went gradually, not significant. This condition cause's noise dissipation process cannot effectively be done on multispectral data, considering the last

MNF band still preserve a lot of information. On the contrary, hyper spectral data may still preserve a great amount of information even after noise band removal due to its many of bands.

We are trying to compare multispectral and hyper spectral image in term of different number of bands. Hyperion as hyper spectral image was chosen in the urban area with sufficient number of vegetation canopy pixels. The Sentinel-2A and Hyperion images processed using MNF transformation to derive the eigenvalue of both images. The derived eigenvalue of both images shown in Table 2. We are selecting the bands used in MNF transformation from 400 – 900 nm wavelength (visible bands), hence Sentinel-2A contain 4 bands and Hyperion contain 46 bands.

Table 2. Comparison of Sentinel-2A and Hyperion Eigenvalue

Sentinel-2A (visible)		Hyperion (visible)	
MNF Band	Information Acc. (%)	MNF Band	Information Acc. (%)
1	71.52	1	28.03
2	87.17	27	83.32
3	95.92	41	95.84
4	100.00	42	96.68
		43	97.52
		44	98.35
		45	99.18
		46	100.00

Analyzing from the Table 2, we found that different number of bands has different accumulation number of information on sequences band number. The smaller number of bands made greater of information losses, otherwise the greater number of bands, smaller information losses. We presume the noise contain in the last bands of each image. The yellow cells of Sentinel-2A and Hyperion MNF band in Table 2 may be the accumulation of each images noise. As for Sentinel-2A MNF band 4 contain 4.08% information losses, leaving the 95.92% of better-quality image. Hyperion MNF band 44, 45, 46 presume as noise band contain 2.48% of information losses, leaving 97.52% of better-quality image. Even though Sentinel-2A removing a smaller number of bands, in reality the information losses are greater than the Hyperion image which removed of its 3 bands.

The analysis resulting the application of MNF for multispectral images with limited number of bands is not effective than application to hyper spectral image due to the process to discard some information aside from the fixed noise elimination process. Statistically, the removal of Sentinel-2A MNF band 4 in this study influenced the vegetation density model accuracy. However, the effect on accuracy is insignificant given the difference of accuracy with and without MNF is only 0.513%.

5. CONCLUSION

The vegetation canopy density modeling accuracy on the image through the MNF process reached 90.889%, lower than the modeling image accuracy without MNF that reached 91.402%. The limited number of bands forced some information losses alongside the noise and unfavorable when compared to hyper spectral image even though less in removed band number. The insignificant changes in noise-information degree causing unsightly change on the canopy density modeling, statistically and visually.

6. REFERENCES

Bhandari, A., Kumar, A., & Singh, G. (2012). Feature Extraction Using Normalized Difference Vegetation Index (NDVI): A Case Study of Jabalpur City, Procedia Technology, 6(2012) pp 612-621 [Crossref]

Boardman, J. W., & Kruse, F. A., 1994, Automated Spectral Analysis: A Geological Example Using AVIRIS Data, North Grapevine Mountains, Nevada: in *Proceedings, ERIM Tenth Thematic Conference on Geologic Remote Sensing, Environmental Research* Institute of Michigan, Ann Arbor, MI, pp. I-407 - I-418.

Chen, L., Yang, X., & Zhen, G. (2017). Potential of Sentinel-2 Data for Alteration Extraction in Coal-bed Methane Reservoirs. *Ore Geology Reviews*, (2017) [Crossref]

Danoedoro, P. (2012). *Pengantar Penginderaan Jauh Digital*. Yogyakarta: Andi Offset.

Ebadi, L., Shafri, H.M., Mansor, S., & Ashurov, R. (2013). A Review of Applying Second-generation Wavelets for Noise Removal from Remote Sensing Data. *Environmental Earth Sciences*, *70*(6), pp 2679–2690 [Crossref]

Exelis. (2015). *Background: MNF Transform*. In part of ENVI Help by Exelis Visual Information Solutions, Inc. a subsidiary of Harris Corporation. USA.

Fauzan, M.A., Kumara, I.S.W., Yogyantoro, R.N., …., & Wicaksono, P. (2017). Assessing the Capability of Sentinel-2 Data for Mapping Seagrass Percent Cover in Jerowaru, East Lombok*. Indonesian Journal of Geography*, *49*(1) pp. 121 – 134. [Crossref]

Frassy, F., Via, G., Maianti, P., & Gianinetto, M. (2013). Minimum Noise Fraction Transform for Improving the Classification of Airborne Hyperspectral Data: Two Case Studies. In *5th Workshop on Hyperspectral Image and Signal Proceeding: Evolution in Remote Sensing.* June 2013. [Crossref]

Green, A. A., Berman, M., Switzer, P., & Craig, M.D. (1988). A Transformation for Ordering Multispectral Data in Terms of Image Quality with Implications for Noise Removal. *IEEE Transactions on Geoscience and Remote Sensing*, *26*(1), pp 65–74 [Crossref]

Jia, X.P., Kuo, B., & Crawford, M.M. (2013). Feature Mining for Hyperspectral Image Classification. *Proc. IEEE Transactions on Geoscience and Remote Sensing*, *101*(3), pp 676–697 [Crossref]

Lantzanakis, G., Mitraka, Z., & Chrysoulakis, N. (2016). Comparison of Physically & Image Based Atmospheric Correction Methods for Sentinel-2 Satellite Imagery. In *Proc. of SPIE Vol. 9688, Fourth International Conference on Remote Sensing and Geoinformation of the Environment (RSCy2016)* [Crossref]

Letexier, D., & Bourennane, S. (2008). Noise Removal from Hyperspectral Images by Multidimensional Filtering. *IEEE Transactions on Geoscience and Remote Sensing*, *46*(7), pp 2061–2069. [Crossref]

Luo, G., Chen, G., Tian, L., Qin, K. & Qian, S. (2016). Minimum Noise Fraction versus Principal Component Analysis as a Preprocessing Step for Hyperspectral Imagery Denoising, *Canadian Journal of Remote Sensing*, *42*(2), pp 106-116. [Crossref]

Navarro, G., Caballero, I., Silva, G., Parra, P., Vazquez, A., & Caldeira, R. (2017). Evaluation of Forest Fire on Madeira Island Using Sentinel-2A MSI Imagery. Int. J. Appl. Earth Obs. *Geoinformation, 58*(2017) pp 97-106 [Crossref]

Pesaresi, M., Corbane, C., Julea, A., Florczyk, A., Syrris, V., & Soille, P. (2016). *Assessment of the added-value of Sentinel-2 for detecting built-up areas. Remote Sensing, 8*(4), 299. [Crossref]

Rouse, J.W., Hass, R.H., Schell, J.A., & Deering, D.W. (1974). Monitoring Vegetation Systems in the Great Plains with ERTS. In *Proceedings of Third Earth Resources Technology Satellite-1 Symposium* SP-351, pp 3010-3017

Traganos, D. & Reinartz, P. (2017). Mapping Mediterranean Seagrasses with Sentinel-2 Imagery, *Marine Pollution Bulletin*, (2017) [Crossref]