A PSNR Review of ESTARFM Cloud Removal Method with Sentinel 2 and Landsat 8 Combination

Dietrich G. P. Tarigan, Sani M. Isa

Computer Science Department
BINUS Graduate Program-Master of Computer Science, Jakarta, Indonesia

Abstract—Remote sensing images with high spatial and temporal resolution (HSHT) for GIS land use monitoring are crucial data sources. When trying to get HSHT resolution images, cloud cover is a typical problem. The effects of cloud cover reduction using the ESTARFM, one of spatiotemporal image fusion technique, is examined in this study. By merging two satellite photos of low-resolution and medium-resolution images, the Enhanced Spatial and Temporal Adaptive Reflectance Fusion Method (ESTARFM), predicts the reflectance value of the cloud cover region. ESTARFM, on the other hand, employs both medium and high-resolution satellite pictures in this study. Using Sentinel 2 and Landsat 8, the Peak Signal Noise Ratio (PSNR) statistical methods are then utilized to evaluate the ESTARFM. The PSNR explain ESTARFM cloud removal performance by comparing the level of similarity of the reference image with the reconstructed image. In remote sensing, this hypothesis was established to get high-quality HSHT pictures. Based on this study, Landsat 8 images that have been cloud removed with ESTARFM may be classed as good. The PSNR value of 21.8 to 26 backs this up, and the ESTARFM result seems good on visual examination.

Keywords—*Cloud removal; RS-Remote sensing; PSNR-Peak signal noise ratio; GIS-Geographic information system; spatiotemporal image fusion*

I. Introduction

An effective way to monitor land-use changes on the earth's surface is remote sensing technology [1]. However, there is still a significant gap between the obtained image and the required image. A grave challenge to maximize the usage of small sensing images (RS) is the cloud problem. Satellite sensors during the acquisition process are unable to capture passive radiation energy in cloud-covered areas, resulting in missed information in satellite images [2]. Electromagnetic waves from ground features are blocked from reaching the sensor system by the cloud. It substantially limits knowledge gained from optical RS images, thus putting a halt to future analysis. Since a result, cloud removal has become a hot issue in the RS community, as it expands the applications and research that optical RS data may be used in Geographic Information System (GIS).

The ESTARFM method is one of many cloud removal methods in remote sensing. The improved spatial and temporal adaptive reflectance fusion model (ESTARFM), which forecasts cloud-covered regions by fusing two satellite pictures, is abbreviated as this model [3]. ESTARFM requires two sets of low and high spatial resolution images captured on two different days (t1 and t2), as well as one collection of images for cloud removal (tp). In the main paper, ESTARFM using Landsat 7 and MODIS imagery. For this research, authors use Sentinel 2 imagery combined with Landsat 8 imagery to run the ESTARFM doing image fusion to correct the cloud effect. The combination of Landsat 8 and Sentinel 2 is chosen to maximize the cloud-covered area result of the ESTARFM method, especially in the area of interest Jakarta, Indonesia. This option combination needs to be investigated because Sentinel 2 imagery is high-resolution imagery, and Landsat 8 is low-resolution imagery. This hypothesis is then evaluated by comparing the ESTARFM result imagery to the actual Landsat 8 imagery image.

Evaluating ESTARFM performance commonly uses statistical analysis like RMSE or MSE [3]. PSNR is another statistical analysis that is possibly used. It is an objective evaluation parameter comparing a reconstructed image to its original image [5]. Besides the final quantitative analysis of PSNR, this paper would also like to visually analyze the influence of Sentinel imagery, cloud density, and specific parameter used. This paper consists of several sections. In Section 2, we presented some current methods used in cloud removal. Section 3, we introduce the ESTARFM method. Then Section 4 explains the research methodology. Section 5 contains the results and experiments. Finally, in Section 6, we bring the paper to a close.

II. Research History

Several prior research has looked towards cloud removal. Cloud removal in remote sensing is the process of reconstructing lost information [4]. According to the previous study, cloud removal methods may be divided into four categories: spatial-based, spectral-based, temporal-based, and multi-source-based [5]. These four approaches are explained below.

The primary purpose of the spatial-based cloud removal model is to provide a complimentary area of cloud covered by exploiting no cloud area of the target image. It estimates pixel loss by spatial interpolation approach as the first type method [6, 7, 8]. Unfortunately, the spatial cloud removal model only performs well on small initial cover gaps and poorly for extensive cloud contamination. So, spatial-based cloud removal can only produce good cloud-free visualization for small holes but is less suitable for quantitative analysis.

Spectral-based cloud removal model can use the correlation of auxiliary clear band and cloud contaminated band to reconstruct the new cloud-free imagery [9]. One of many

approaches is a fog-optimized transformation method [10], correcting the visible radiometric bands from contamination of thin clouds and haze in Landsat imagery. This method is known as the HOT (haze optimization transformation) approach. The HOT change is a visual band space analysis used to see various classes of land use cover. Unfortunately, the spectral response of fog in the visible band space is extremely sensitive to the fog's optical wavelength and depth. The reconstruction result of this approach is quite good in a thin layer of moisture and cannot overcome thick cloud cover. In addition, the spectral-based cloud removal method is unable to distinguish sure land covers such as ice/snow, terrain, and water bodies.

Multitemporal imagery for a given place can be produced by RS systems with frequent satellite revisit cycles (n-day cycles). The temporal cloud removal approach fully uses the temporal correlation between multitemporal images for reconstruction [11, 12]. The material method can effectively reconstruct the lost area due to cloud contamination, especially in congested acquisition frequency of time series data. This method follows solid time series displaying common cloud pollutants when the images are chronologically ordered [4]. Unfortunately, the temporal process is limited to congested data sets, often provided by coarse spatial resolution imagery. On the other hand, the material approach implies that spatial land coverage is determined during the acquisition interval [5]. Temporal-based techniques are appropriate for den situations, according to this general idea.

Combining the advantages of the above methods can solve several problems in cloud removal methods. This approach is called the multisource-based or hybrid approach. The auxiliary picture must have the same wavelength and spatial resolution as the target image in this method [13]. The spatiotemporal method of cloud removal is an example of a hybrid approach that Zhu [14] has been developed. The ESTARFM Model is a commonly used benchmark for developing cloud removal models that employ the spatiotemporal fusion approach. [13]. This algorithm was developed based on STARFM as an improvement.

III. ESTARFM METHOD

Various satellite sensors acquired on the same day and in the exact location can be compared and correlated, mainly when preprocessing steps such as radiometric correction, geometric correction, and atmospheric adjustment have been completed. However, there can be a systematic bias in surface reflection across various sensor images due to sensor system variations such as orbital parameters, bandwidth, acquisition time, and spectral response function. This condition was the primary purpose of ESTARFM to utilize the correlation of various satellite sensor data. Then, integrate multi-source data while minimizing system bias ESTARFM looking at the heterogeneity of the land surface, pure pixels, and mixed pixels [14]. Images with low spatial resolution but high temporal resolution are referred to as "coarse-resolution" by ESTARFM.

In contrast, images with high spatial resolution but low temporal resolution are referred to as "fine resolution" images. [13]. Eq. 1 examines the first steps and equations for spectrally comparable homogeneous pixels within a moving window.

$$\left| L(x_i, y_i, t_k) - L\left(x_{\frac{w}{2}}, y_{\frac{w}{2}}, t_k\right) \right| \leq 2 \times \frac{\sigma(B)}{n} \qquad (1)$$

where, $L(x_i, y_i, t_k)$ represent the i-th pixels surface reflectance with the location (x_i, y_i) in the moving window on the observed data t_k. $L\left(x_{\frac{w}{2}}, y_{\frac{w}{2}}, t_k\right)$ is the surface reflectance position of the center pixel $\left(x_{\frac{w}{2}}, y_{\frac{w}{2}}\right)$ in the moving window on the observed data t_k [13]. $\sigma(B)$ displays the whole image's standard deviations for band B.; n shows how many classes are in the study area. A homogeneous pixel is spectrally similar if it fulfills the conditions in equation (1). After a homogeneous pixel is obtained, the weight and conversion coefficient is calculated. Equation 2 is then used to estimate the expected reflection of a center pixel in a moving window.

$$F\left(x_{\frac{w}{2}}, y_{\frac{w}{2}}, t_p, B\right) = F\left(x_{\frac{w}{2}}, y_{\frac{w}{2}}, t_0, B\right) + \sum_{i=1}^{N} W_i \times V_i \times$$
$$(C(x_i, y_i, t_p, B) - C(x_i, y_i, t_0, B)) \qquad (2)$$

where N is the quantity of spectrally uniform pixels with similar spectral characteristics, including the moving window's center pixel; (x_i, y_i) indicates the pixel location similar to i-th; w indicates the width of the search window; t_0 and t_p are the observation date and the predicted date; W shows the corresponding weight and V_i is conversion coefficient of the i-similar pixel; $F\left(x_{\frac{w}{2}}, y_{\frac{w}{2}}, t_p, B\right)$, $F\left(x_{\frac{w}{2}}, y_{\frac{w}{2}}, t_0, B\right)$ indicate the reflectance of acceptable spatial resolution of the center pixel on the predicted and observed dates, respectively; $C(x_i, y_i, t_p, B)$, $C(x_i, y_i, t_0, B)$ represent the coarse spatial resolution reflectance of the pixels located at (x_i, y_i) at the predicted and observed data [13]. The conversion coefficients are calculated using a linear regression model into a similar pixel of high and low spatial resolution imagery [15]. W_i is the weight for similar pixels is calculated based on the spatial distance from the middle pixel of the moving window and the spectral similarity between fine and coarse resolution pixels [13]. W_i is calculated by equation.

$$d_i = 1 + \frac{\sqrt{(x_i - x_{w/2})^2 + (y_i - y_{w/2})^2}}{\frac{w}{2}} \qquad (3)$$

$$D_i = (1 - R_i) \times d_i \qquad (4)$$

$$W_i = \frac{1/D_i}{\sum_{i=1}^{N} 1/D_i} \qquad (5)$$

where, d_i denotes the geographic distance between the center and a spectrally similar uniform pixel in a moving window; R_i denotes the spectral correlation coefficient calculated using the correlation between any similar pixels in a fine spatial-resolution image and a corresponding coarse spatial-resolution pixel; and D_i is an index integrating spectral and spatial similarity [13]. ESTARFM additionally considers the temporal weight when predicting the center pixel's reflection. The following equation is used to determine the final projected reflectance.

$$F\left(x_{\frac{w}{2}}, y_{\frac{w}{2}}, t_p, B\right) = T_m \times F_m\left(x_{\frac{w}{2}}, y_{\frac{w}{2}}, t_p, B\right) + T_n \times$$
$$F_n\left(x_{\frac{w}{2}}, y_{\frac{w}{2}}, t_p, B\right) \qquad (6)$$

where, T_k $(k \in (m, n))$ shows the temporal weight between observed and estimated dates.

IV. RESEARCH METHODOLOGY

This research consists of four stages, namely: the data collection stage, preprocessing step, the implementation stage, and the evaluation stage. At the data collection stage, the location is the province of Jakarta and timespan around 2020. The images to be used are Sentinel 2 and Landsat 8. The reason for Sentinel 2 images is chosen because its 10-day temporal resolution is expected to be able as an auxiliary image in the spatial, and the spectral resolution of Sentinel 2, which reaches 10 meters, is expected to improve cloud-covered result data of ESTARFM.

In preprocessing stage, the initial steps are done to Sentinel 2 and Landsat 8 imagery. Radiometric, atmospheric, and geometric corrections in the preprocessing setting are performed on SNAP (Sentinel Application Platform) software for Sentinel 2 images and SCP (Semi-Automatic Classification Plugin) plugin [16] for Landsat 8 images in QGIS. Specifically for Sentinel 2 images, an up-sampling process [17] was carried out to equalize the spatial resolution with Landsat 8.

Pairs of fine and coarse resolution images taken on the neighboring date and a set of rough resolution images for the target forecast date are required to execute the ESTARFM model [16]. Fig. 1 states the process flow of the ESTARFM image fusion. Before applying ESTARFM, all images must be processed first to reflect the geo-registered surface.

When it comes to putting ESTARFM into action, there are four key stages [14]. To begin, two fine-resolution and coarse-resolution imagery search for pixels in the local window that is similar to the center pixel. Second, the weights (Wi) of all matched pixels are computed. Then, two pairs of fine and coarse images in the designated frame are used separately to find similar pixels. This step can be seen in Fig. 2 (A). Third, linear regression is used to calculate the conversion coefficient Vi of the fine and coarse images. The conversion coefficient conforms to the yield obtained using regression analysis of identified similar pixels. This step can be seen in Fig. 2 (B). On the target date, the reflectance value of each pair of input imagery is collected. This step can be seen at Fig. 3. So, at the appropriate forecast date, Wi and Vi are utilized to compute the fine-resolution reflection of the coarse-resolution imagery [13]. The weighted average is used in the fourth step, which takes into account three factors: the spectral difference between fine and coarse images taken on the same date, the temporal difference between the images, and the spatial distance between similar pixels in the moving window and the center pixel [14].
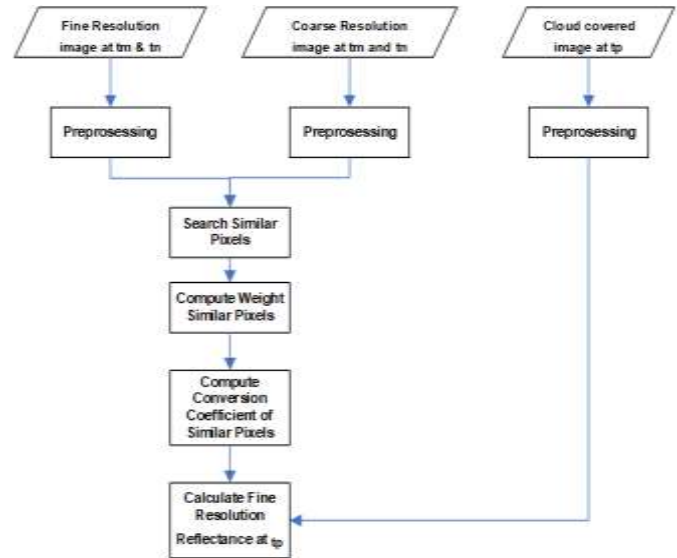

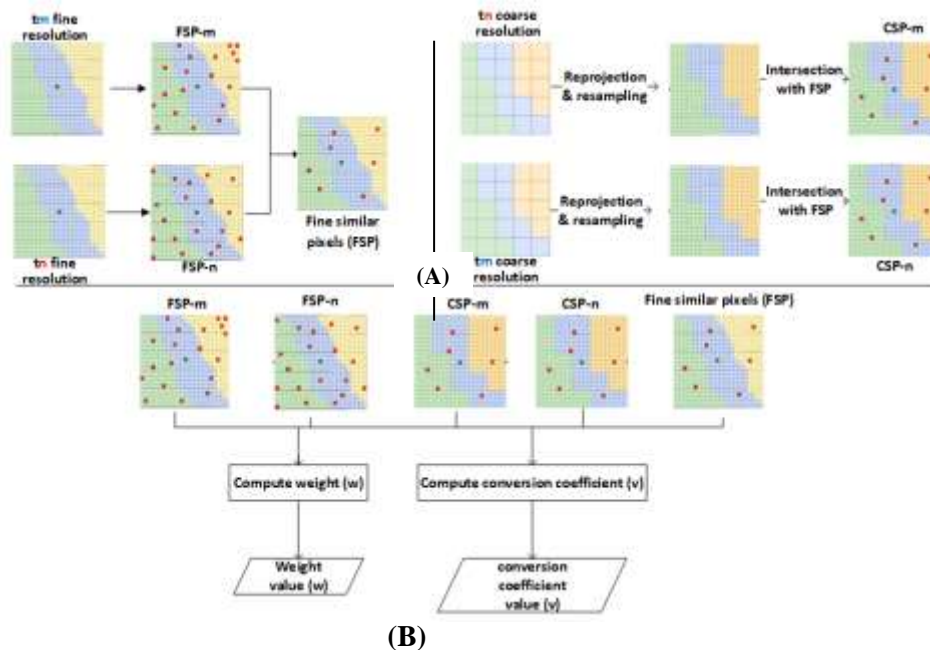Fig. 1. ESTARFM Image Fusion Method.

(A)

(B)

Fig. 2. (A) - Identify Similar Pixels Process; (B) - The Process of obtaining Conversion Coefficient.
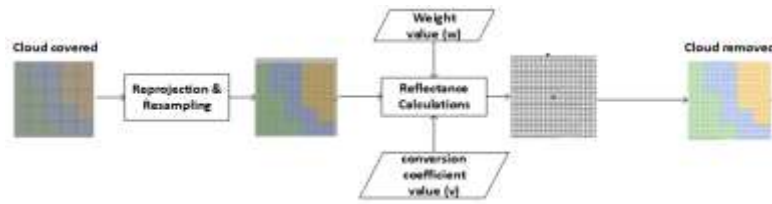
Fig. 3.   Reflectance Calculation Process of Cloud-covered Area.

The requirement of two pairs of input images limits the possible combinations of inputs [13]. Also, because of the long processing times for these methods, only one input date combination per target is tested. The weighted average technique obtains the fine image on the date the Landsat 8 image is cloud-covered. Similar pixels play the primary role in the performance of weighted-function-based spatiotemporal fusion methods [18]. This study quantitatively assessed the performance of ESTARFM from the perspective of similar pixels.

ESTARFM also studies influencing factors of similar-pixel identification like the methods of similar-pixel identification, the number of classes (m), and moving window size (w) [13]. Enlarging the period between the base and forecast dates additionally assesses the effect of input base picture pairings. This study could provide a method to analyze the performance of weighted-function-based methods and guidance for the future applications of ESTARFM in a specific area. This entire cycle has been done using a python-based algorithm taken from Zhu's web page in https://xiaolinzhu.weebly.com/.

Following that descriptive/qualitative and analytical/quantitative assessments of the fusion product quality are done when the cloud-free image results have been acquired. The qualitative analysis focuses on a visual examination of variations in reflection and spatial patterns in observed (actual) and expected (cloud-free) pictures [19, 20]. Meanwhile, the PSNR is used for quantitative evaluation. The "Peak Signal-to-Noise Ratio" (PSNR) is a statistical technique used in remote sensing applications to assess data quality following picture fusion, mainly from various satellite sensors [21]. The PSNR value may be used to evaluate the amount of similarity among the original image and the reconstructed/predicted image after fusion to determine the quality of the reconstructed image. The PSNR is calculated using the equation below.

$$PSNR = 20 \times log10 \frac{N}{\sqrt{MSE}} \qquad (7)$$

and

$$MSE = \frac{1}{mn}\sum_{y=1}^{m}\sum_{x=1}^{n}(I(x,y) - I'(x,y))^2 \qquad (8)$$

N is the most significant potential pixel value in a picture, I (x, y) is the pixel value difference between the original and predicted image. At the same time, m and n are the image's length and breadth, respectively. PSNR is a compression parameter that is often utilized [22]. The amount of resemblance between the compressed or reconstructed picture and the original image is considered this parameter. The mean squared error (MSE) of the two pictures being compared determines the PSNR value. This PSNR value is comparable to

the MSE value. Thus, if the MSE between the compressed image and the slice image is less, the PSNR value between the two images will be higher [23]. This paper assumes noise as a cloud-covered area, so a satellite sensor can't receive a signal. The peak ratio between noise and signal also represents the performance of ESTARFM, especially when use Landsat 8 and Sentinel 2 combination in Jakarta province during 2020.

V.   EXPERIMENTS AND DISCUSSION

A. Data

The research area defined in this paper is Jakarta, Indonesia, as seen in Fig. 4. This area was chosen because of the frequent occurrence of cloud cover, and the site is relatively small, around 661.5 km^2. Jakarta is determined to reduce mistakes of image fusion due to land-use change. Because Jakarta, as a metropolitan city, is considered to have insignificant land-use changes. In Landsat 8 imagery, Jakarta is covered in tile path/row 122/64 and Sentinel 2 in two tile paths/rows, namely 49 MXU and 49 MYU. The technical specifications of these two images can be seen in Table I. The Sentinel 2 image is chosen as a fine-resolution image because of its five days temporal resolution, compared to Landsat 8 image, which has 16 days temporal resolution and is considered coarse resolution. The smaller the temporal resolution, the higher the chance of getting an image without cloud cover. In addition, Sentinel 2 has a spatial resolution of 10m on the RGB band compared to the 30m that Landsat 8 has on the same band. The more detailed spatial resolution is expected to increase the ESTARFM model's ability to correct cloud-covered Landsat 8 images.
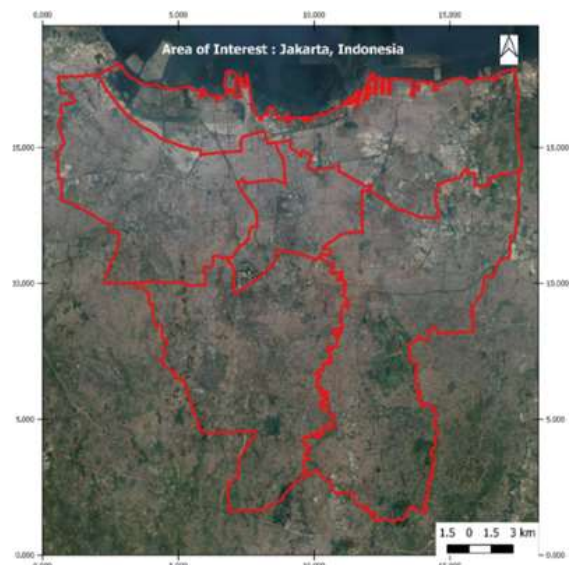

Fig. 4.   Area of Interest: Jakarta, Indonesia.

TABLE I.   THE FEATURES OF THE LANDSAT-8 AND SENTINEL-2 DATASETS USED IN THIS STUDY

Characteristic	Landsat-8 OLI	Sentinel-2 MSI
Bands (spectral range) used	B2 (0.452 - 0.512 μm)	B2 (0.439 - 0.535 μm)
	B3 (0.533 - 0.590 μm)	B3 (0.537 - 0.582 μm)
	B4 (0.636 - 0.673 μm)	B4 (0.646 - 0.685 μm)
Spatial Resolution	30m	10m
Temporal Resolution	16 days	5 days
Research Location Tile	WRS Path/Row: 122/64	Path/Row: 49MXU 49 MYU

The Sentinel 2 and Landsat 8 satellite imagery data used are the 2020 acquisitions year. Direct visual checks are carried out on the web https://glovis.usgs.gov to select satellite images with certain clouds cover manually. According to the findings, obtaining satellite imagery that is cloud-free by 2020 would be challenging. Correcting and identifying cloud-covered areas in Landsat 8 data, a cloud-free picture is selected from Sentinel 2 and Landsat 8 satellite images with a cloud cover percentage of 0% to 5%.

Several images were used on three simulations, as shown in Table II and Table III. Visually, the condition of Landsat 8 and Sentinel 2 data can be seen in Fig. 5 and the target cloud removal in Fig. 6. As long as in 2020, based on the author's visual analysis via the web https://glovis.usgs.gov, these two Landsat 8 and Sentinel 2 images are considered cloud-free, so they can be used as correcting Landsat 8 photos in the ESTARFM model.

All Landsat 8 and Sentinel 2 data are preprocessed first. Landsat 8 imagery is done in the Semi-Automatic Classification Plugin (SCP) plugin in QGIS, and Sentinel 2 is done in SNAP. Landsat 8 imagery that has been preprocessed, such as radiometric correction and reflectance conversion, is then cropped based on Jakarta's area of interest. After that, reprojection is carried out to the UTM WGS84 48S coordinates according to the location of the Jakarta area.

This treatment is carried out on Band 2 3 4 of Landsat 8 and combined into one complete RGB image on each date. The Sentinel 2 image is preprocessed in SNAP for radiometric correcting and up-sampling the resolution from 10m to 30m. Because Jakarta is located on two tiles, Sentinel 2, namely, 48MXU and 48MYU, mosaicking is done on the SNAP program to combine two tile images on each date. Furthermore, after the two tiles are combined, they are cropped based on Jakarta's area of interest. This treatment is also carried out on band 2 3 4 Sentinel 2 and then combined into one complete RGB image on each date.

B. Implementation Details

ESTARFM [17] python source code has been provided on Zhu's web page. This research uses the PyCharm IDE platform to run the python code. PyCharm is a free and open-source integrated development environment (IDE) for the Python programming language. JetBrains, a Czech Republic-based firm, created this program.

It should be noted that the parameter set in the ESTARFM model needs to be determined first due to Jakarta, as the area of research is different from the original ESTARFM model. The preliminary trial is done to determine the optimal value to be used for the Jakarta area. After parameters are set, the ESTARFM model can be run immediately.

When running, the writer simply points to the desired GeoTIFF file image directory. After the process is finished, a cloud-removed image is generated with the GeoTIFF file extension.

In addition, QGIS is used to extract the reflectance value of the reference image and the model result image. The extraction procedure begins with creating a point-shaped shapefile spread out throughout the Jakarta region of interest. About 10 000 scattered points are assumed as a ground check for the reflectance value of the image. The reflectance values obtained are for RGB or band 2 3 4. Each reflectance value is processed in a spreadsheet to make a scatterplot and PSNR analysis based on the ESTARFM method section.

TABLE II.   THE LANDSAT-8 AND SENTINEL-2 REFLECTANCE PRODUCTS USED AS THE BASE IMAGERY

Satellite Imagery	Product ID	Date Acquisition	Remarks
Landsat-8 OLI	LC08_L1TP_122064_20200422_20200508_01_T1	April 22, 2020	Coarse (C) t1
	LC08_L1TP_122064_20200913_20200920_01_T1	September 13, 2020	Coarse (C) t2
Sentinel-2 MSI	S2A_MSIL2A_20200728T025551_T48MXU_20210213T103728.SAFE	July 28, 2020	Fine (F) t1 - stacked image
	S2A_MSIL2A_20200728T025551_T48MYU_20210213T103728.SAFE		
	S2A_MSIL2A_20200827T025551_T48MXU_20210213T111615.SAFE	August 27, 2020	Fine (F) t2 - stacked image
	S2A_MSIL2A_20200827T025551_T48MYU_20210213T111615.SAFE		

TABLE III.   THE LANDSAT-8 CLOUD-COVERED TARGET IMAGE

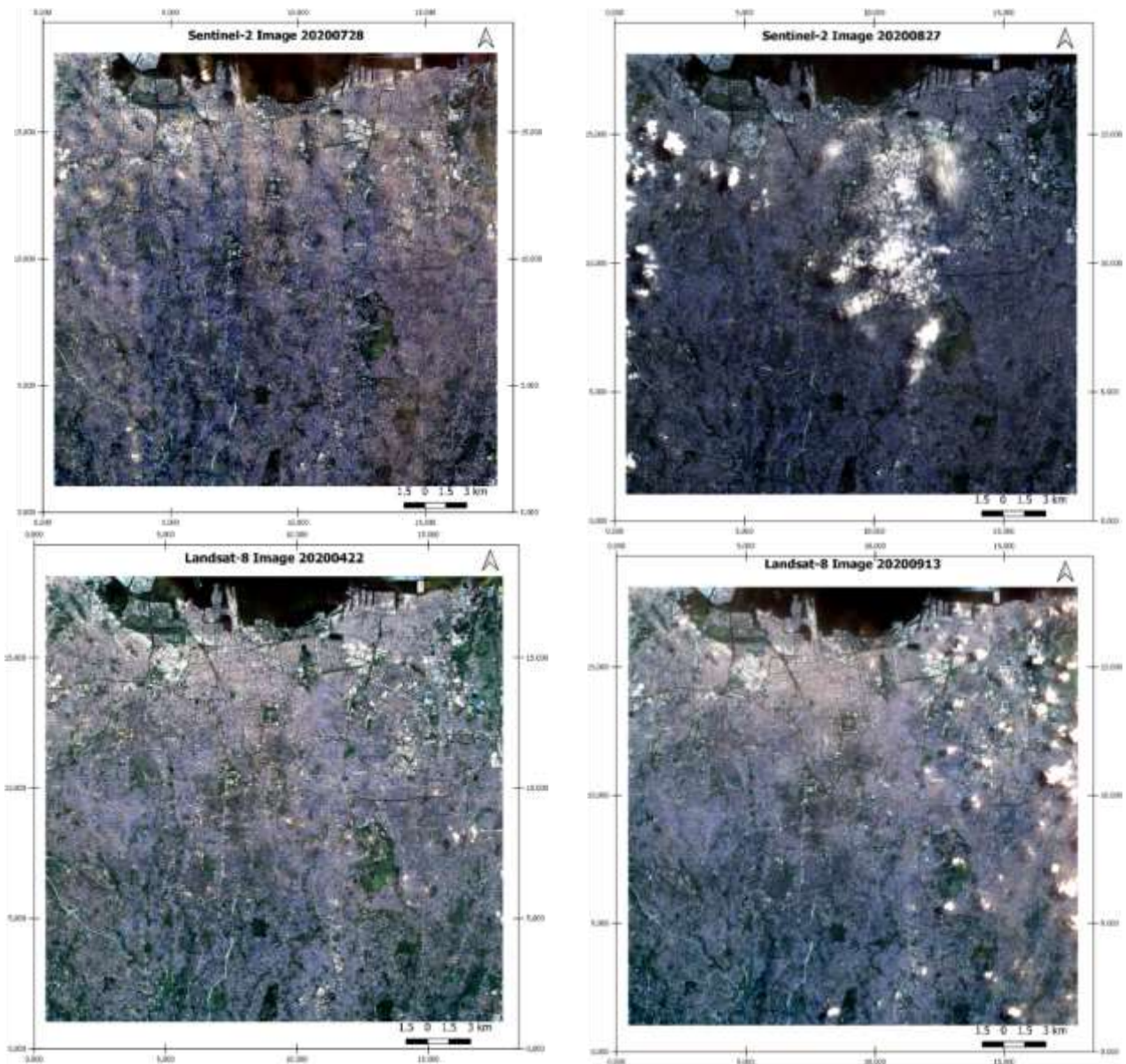| Satellite Imagery	Product ID	Date Acquisition	Remarks
Landsat-8 OLI	LC08_L1TP_122064_20200625_20200707_01_T1	June 25, 2020	Cloud covered (tp) -sim1
	LC08_L1TP_122064_20201015_20201104_01_T1	October 15, 2020	Cloud covered (tp) -sim2
	LC08_L1TP_122064_20201202_20201217_01_T1	December 2, 2020	Cloud covered (tp) -sim3

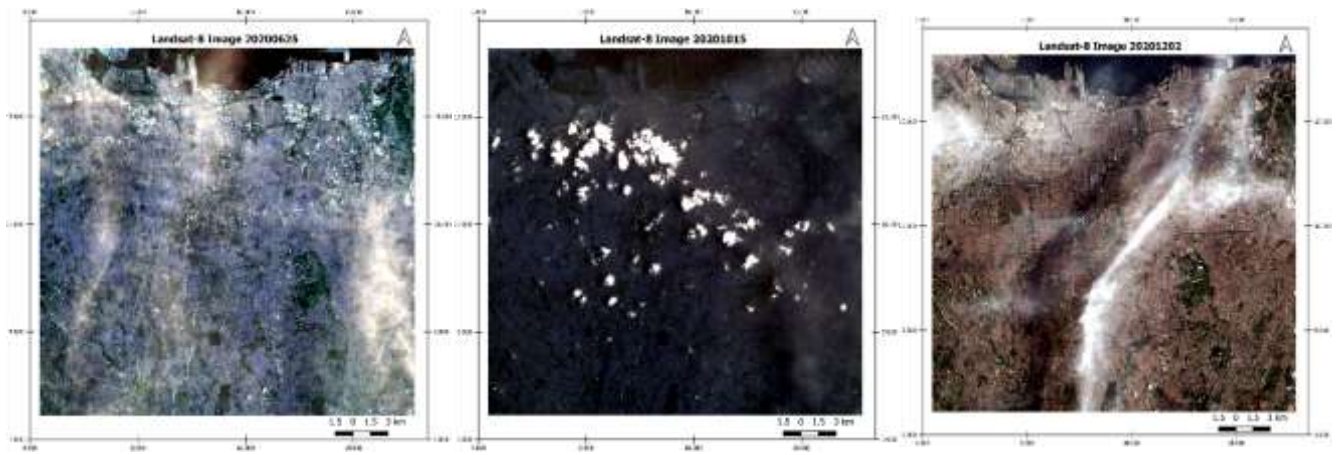Fig. 5.   Landsat 8 and Sentinel 2 relatively Cloud Free Image as based Image.

Fig. 6.   Landsat 8 Target Image for Cloud removal.

TABLE IV. ESTARFM PARAMETER USED FOR A STUDY AREA

Parameter	Value	Remarks
w	25	# set the half window size. If 25, the window size is 25*2+1=51 fine pixels
num_class	3	# set the estimated number of classes, please set a larger value if blending images with very few bands
DN_min	0	# set the range of DN value of the image,If byte, 0 and 255
DN_max	2	
background	-9999	# set the value of background pixels. 0 means that pixels will be considered as background if one of its bands= 0
patch_long	500	# set the size of each block,if process whole ETM scene, set 500-1000

C. Discussion

At the initial stage of data acquisition, the difficulties in obtaining an image that meets the research needs. The research condition is determined that the correcting image uses 0% to 5% cloud cover percentage. This is because of Jakarta area is quite often covered in clouds. Besides, the resolution of both Sentinel 2 and Landsat 8 imagery in RGB band is 10m, and 30 m and Jakarta area are 662.5 km^2, if the cloud is covered, it will almost cover all the images. In addition, the Landsat 8 16-day temporal resolution is quite challenging to get cloud-free pictures. Even though the spatial resolution is just 10m and is sometimes obscured by clouds, Sentinel 2 imagery is quite helpful in this regard. Still, there are more image options available because the temporal resolution is just five days.

The picture utilized by the ESTARFM model is RGB, but because it uses a GeoTIFF file, the model can see the band-specific data. Landsat-8 and Sentinel-2 spectral bands (bands 2, 3, and 4) are utilized to create synthetic Landsat-like pictures [13]. Because the research region is less diverse, the class labels in Eq. 1 are adjusted to three. Because the pixel vector's dimension was 3 x 1 (band 2, band three, and band 4 in Landsat 8), the criterion in Eq. (1) had to be passed three times to choose a spectrally comparable homogeneous pixel vector [24]. Table IV shows the exact parameters used in this study. As stated in Eq 2, moving windows pick spectrally identical pixels and compute the weight function and conversion coefficient, followed by Eq 3 to 6 [13].

The usage of images with higher temporal and spatial resolution, such as Sentinel 2, to predict the cloud cover area in Landsat 8 is comparable and exhibits spectral similarities. Visually, the prediction image is relatively able to reduce or even expect the cloud-covered site. In the scatterplot distribution, there is a reasonably good correlation. However, on the 25 June 2020 data, the predicted image has a red spot that looks similar to the cloud cover in the reference Sentinel 2

image. This can also be seen in the results of the ESTARFM on the other two data, on August 27th, 2020, and December 2nd, 2020. Likewise, on the data on December 15th, 2020, the cloud noise in the original data has been reduced. However, since the correcting information still contains clouds, the results of the ESTARFM prediction image are also marginally influenced by shadows.

On the other hand, these results show that the ESTARFM model can perform image fusion between Landsat 8 and Sentinel 2. According to the author, there is a possibility that the test will be conducted under the opposite conditions, namely with Sentinel 2, which has a higher spatial resolution and a tighter temporal resolution. As a result of this research, it can be demonstrated that the ESTARFM model can also be used with a Landsat 8 and Sentinel 2 combination.

The ESTARFM parameter setting also determines a good predictive image result. Several test iterations of the model with the parameters used are also sufficient to reconstruct the cloudy data. The parameters that need to be considered in this ESTARFM node are listed in Table IV. Based on the author's trial for the size of interest in Jakarta, the most influential parameters are DN max and CN min, num-class. DN max DN min value is the reflectance value limit that depends on the satellite image data used. Meanwhile, num-class estimates the number of classes that are very influential in combining several image bands so that these parameters need to be explicitly set as required.

PSNR is a frequently used metric to compare the quality of reconstructed pictures to the original image [23]. The PSNR and MSE values are inversely related. This indicates that the larger the PSNR value of the two pictures, the lower the MSE between the compressed image and the slice image. In lossy compression, the compressed image closely resembles the original. Therefore, a higher PSNR value implies more excellent image quality. Meanwhile, there is no difference between the compressed and original images in lossless compression resulting in a PSNR value of infinity [25].

In this study, it can be seen that the PSNR value for each satellite image band is around 26.4 to 27.5 on the acquisition date of June 25th and December 2nd, 2020. Meanwhile, on the acquisition date of October 15th, 2020, the PSNR value ranges from 21.8 to 22.6. Based on this PSNR value, it can be seen that the correlation is that the higher the PSNR value, the results of the ESTARFM prediction image are also closer to the cloud-covered Landsat image. Indeed, in this study, the authors still have difficulty determining the limit of the PSNR value in remote sensing. This is mainly because PSNR is commonly used for image compression as a giant image storage solution. However, based on PSNR writings, the authors assess PSNR above 25 as good for cloud removal cases with ESTARFM [5;12].

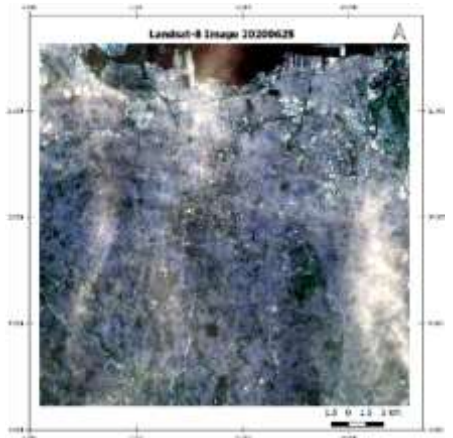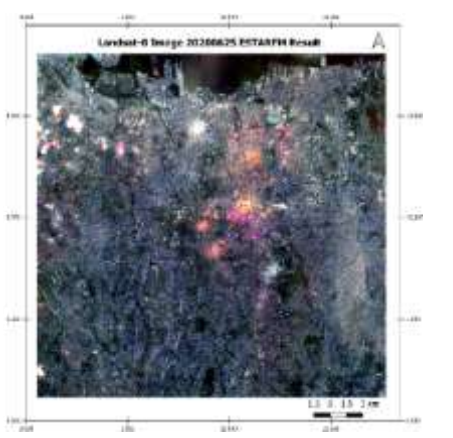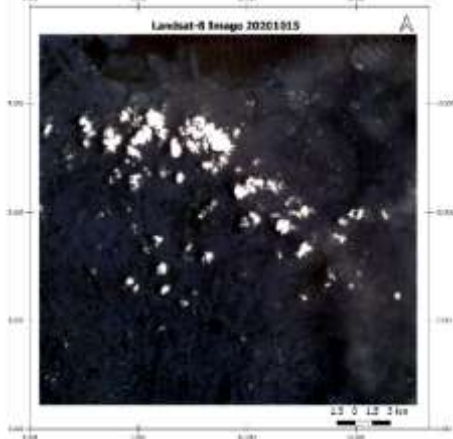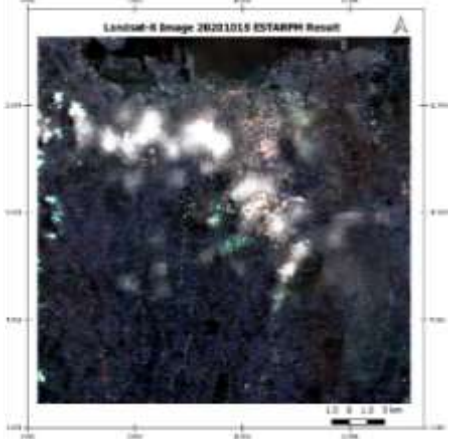TABLE V. VISUAL ASSESSMENT OF OBSERVED AND PREDICTED IMAGE LANDSAT

Date	Observed Images	ESTARFM Predicted Image
June 25th, 2020		
October 15th, 2020		
December 2nd, 2020		

TABLE VI. PSNR PERFORMANCE EVALUATION OF DIFFERENT DATE SIMULATION ON RESPECTIVE BANDS

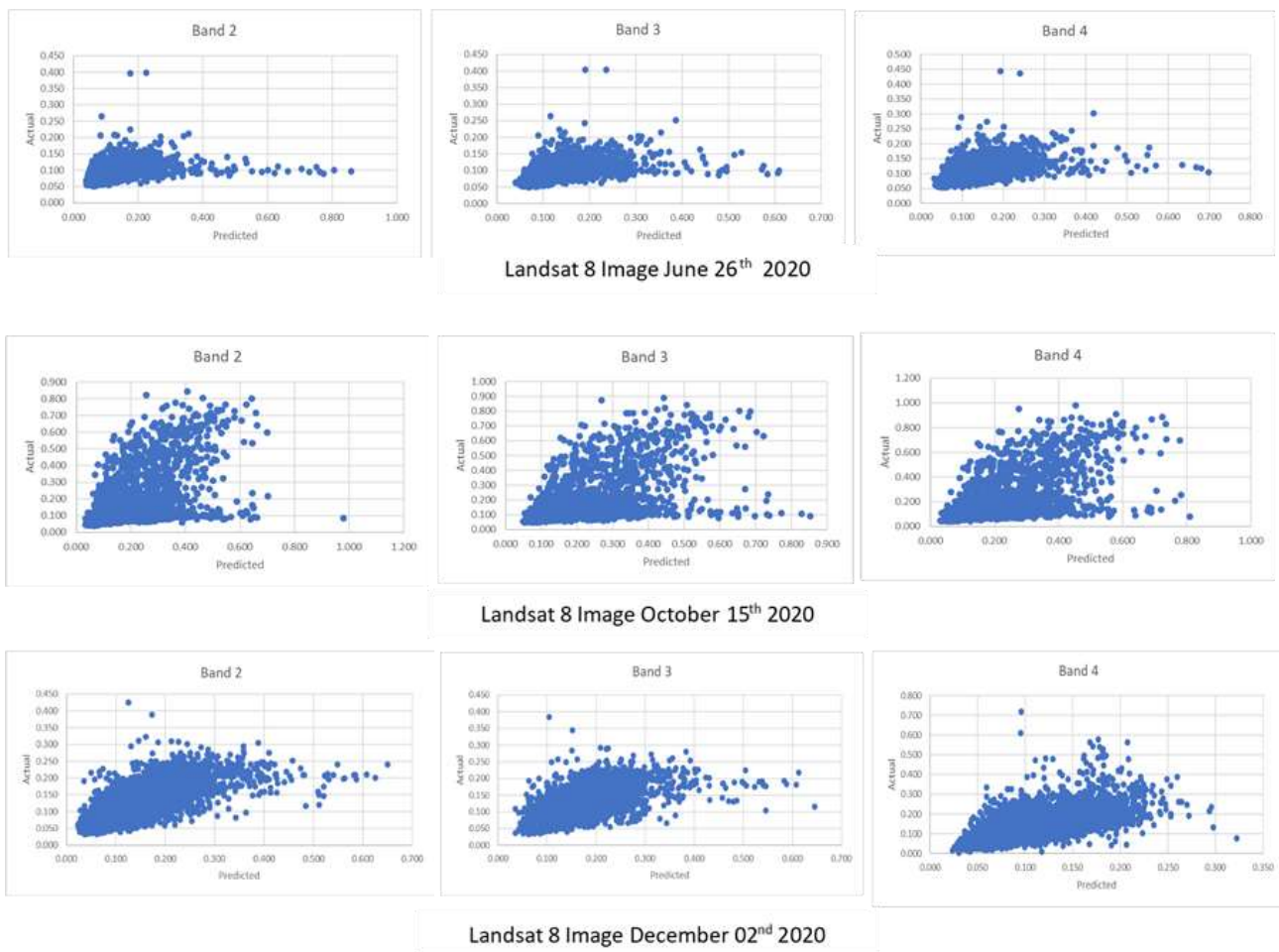| Date	PSNR		
	Band 2 (Blue)	Band 3 (Green)	Band 4 (Red)
June 25, 2020	26.55	27.08	27.42
October 15, 2020	22.61	21.98	21.89
December 2, 2020	27.22	26.42	26.63

Fig. 7. Actual and Projected Landsat Pictures on Corresponding Dates, with Band-specific Density Scatterplots.

In addition, the scatterplot graph, as seen in Fig. 7, can also be seen the correlation between the reflectance value of the cloud-covered image and the predicted ESTARFM image. Overall, the point distribution is close to the diagonal line, but some wrong reflectance values still move away from the trendline. Each scatterplot consists of 10000 points of reflectance values from each band of observation dates. So that the visual analysis of the reflectance value can be done more easily besides seeing the image directly, which cannot be measured and tends to be qualitative, however, the scatterplot analysis can be done quantitatively. So, it can be said that the combination of Landsat 8 and Sentinel 2 in the ESTARFM model for cloud removal has worked well and can correct some cloud-covered values.

VI. CONCLUSION

The primary objective of this research was to contribute a PSNR review of ESTARFM as one of many spatiotemporal fusion frameworks using specific parameters for the area of interest Jakarta. Overall, ESTARFM performs well in our research. The weight-function-based fusion model ESTARFM has successfully blended high spatial-temporal resolution of Sentinel-2 data with the low spatial-temporal resolution of Landsat-8. As seen in Table V, Fig. 2, and Table VI, ESTARFM performance is average beyond good. The further

research possibilities, authors recommended using 100% cloud-free image or below 2% cloud coverage area as based fusion image to predict and correct cloud-covered area of Sentinel-2 and Landsat-8 imagery. One source of satellite imagery that can be used is Lapan. On the official website https://inderaja-catalog.lapan.go.id/ you can find satellite images of various resolutions, especially for the Indonesian region. In addition, further parameter analysis can also be carried out in other regions in Indonesia.

REFERENCES

[1] J. B. Campbell, Introduction to remote sensing, New York: Guilford Press, 1996.

[2] H. Briassoulis, "Analysis of Land Use Change: Theoretical and Modeling Approaches," WVU Research Repository, 2020.

[3] C. Kwan, X. Zhu, F. Gao, B. Chou , D. Perez , J. Li and G. Marchisio, "Assessment of spatiotemporal fusion algorithms for planet and worldview images," Sensors, vol. 18, no. 4, 2018.

[4] H. Shen, X. Li, Q. Cheng, C. Zeng, G. Yang, H. Li and L. Zhang, "Missing information reconstruction of remote sensing data: A technical review," IEEE Geoscience and Remote Sensing Magazine, vol. 3, no. 3, pp. 61-85, 2015.

[5] H. Shen, J. Wu, Q. Cheng , M. Aihemaiti, C. Zhang and Z. Li, "A spatiotemporal fusion based cloud removal method for remote sensing images with land cover changes," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 12, no. 3, pp. 862-874, 2019.

[6] F. Van der Meer, "Remote-sensing image analysis and geostatistics," International Journal of Remote Sensing, 33(18), pp. 5644-5676, 2012.

[7] C. Zhang, W. Li and D. Travis, "Gaps‑fill of SLC‑off Landsat ETM+ satellite image using a geostatistical approach," International Journal of Remote Sensing, vol. 28, no. 22, pp. 5103-5122, 2007.

[8] P. Scaramuzza and J. Barsi, "Landsat 7 scan line corrector-off gap-filled product development," In Proceeding of Pecora, pp. 23-27, 2005.

[9] P. Rakwatin, W. Takeuchi and Y. Yasuoka, "Restoration of Aqua MODIS band 6 using histogram matching and local least squares fitting," IEEE Transactions on Geoscience and Remote Sensing, 47(2), pp. 613-627, 2008.

[10] Y. Zhang , B. Guindon and J. Cihlar, "An image transform to characterize and compensate for spatial variations in thin cloud contamination of Landsat images," Remote Sensing of Environment, vol. 82, no. 2, pp. 173-187, 2002.

[11] C. Lin, K. Lai, Z. Chen and J. Chen, "Patch-based information reconstruction of cloud-contaminated multitemporal images," IEEE Transactions on Geoscience and Remote Sensing, vol. 52, no. 1, pp. 163-174, 2014.

[12] C. Lin, P. Tsai, K. Lai and J. Chen, "Cloud removal from multitemporal satellite images using information cloning," IEEE transactions on geoscience and remote sensing, vol. 51, no. 1, pp. 232-241, 2013.

[13] R. Ghosh, P. Gupta, V. Tolpekin and S. Srivastav, "An enhanced spatiotemporal fusion method – Implications for coal fire monitoring using satellite imagery," International Journal of Applied Earth Observation and Geoinformation, vol. 88, no. 102056, pp. 1-15, 2020.

[14] X. Zhu, J. Chen, F. Gao, X. Chen and J. Masek, "An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions," Remote Sensing of Environment, vol. 114, no. 11, pp. 2610-2623, 2010.

[15] J. Xue, Y. Leung and T. Fung, "An unmixing-based Bayesian model for spatio-temporal satellite image fusion in heterogeneous landscapes," Remote Sens. (Basel) 11 (3), p. 324–346, 2019.

[16] C. Luca, "Semi-Automatic Classification Plugin for QGIS," 2020. [Online]. Available: http://dx.doi.org/10.13140/RG.2.2.25480.65286/1.

[17] J. Louis, R. Debacker, B. Pflug, M. Main-Knorn, J. Bieniarz, U. Mueller-Wilm and F. Gascon, "Sentinel-2 sen2cor: L2a processor for users," in Proceedings Living Planet Symposium 2016, 2016.

[18] K. Hazaymeh and Q. Hassan, "Spatiotemporal image-fusion model for enhancing the temporal resolution of Landsat-8 surface reflectance images using MODIS images," Journal of Applied Remote Sensing,, vol. 9, no. 1, 2015.

[19] J. Ma, W. Zhang, A. Marinoni, L. Gao and B. Zhang, "An improved spatial and temporal reflectance unmixing model to synthesize time series of landsat-like images," Remote Sensing, vol. 10, no. 9, 2018.

[20] J. Ma, W. Zhang, A. Marinoni, L. Gao and B. Zhang, "Performance assessment of ESTARFM with different similar-pixel identification schemes," Journal of Applied Remote Sensing, vol. 12, no. 2, 2018.

[21] L. Bo, Q. Youshan, F. Guilan, Y. Xiufang and X. Bin, "Maximal PSNR wavelet bi-linear interpolation iterative algorithm in remote sensing image," ACTA Photonica Sinica, vol. 35, no. 3, pp. 468-480, 2006.

[22] A. Babaee, S. Shahrtash and A. Najafipour, "Comparing the trustworthiness of signal-to-noise ratio and peak signal-to-noise ratio in processing noisy partial discharge signals," IET Science Measurement & Technology, pp. 112-118, 2013.

[23] N. Kristianti, N. Purnawanti and B. Rolando, "Analisis Pengaruh Citra Gelap, Normal, Terang Terhadap Wavelet Orthogonal.," Jurnal Buana Informatika, pp. 93-100, 2018.

[24] F. Chen, L. Tang and Q. Qiu, "Exploitation of CBERS-02B as auxiliary data in recovering the Landsat7 ETM+ SLC-off image," in 18th International Conference on Geoinformatics, IEEE, 2010.

[25] A. Widipaminto, A. Indradjad, D. Monica and R. Rokhmatuloh, "Analisis Metode Kompresi Berdomain Wavelet Pada Citra Satelit Resolusi Sangat Tinggi," Jurnal Penginderaan Jauh dan Pengolahan Data Citra Digital, vol. 16, no. 1, 2020.