

ASSESSMENT OF DIFFERENT IMAGE TRANSFORMATION METHODS ON DIWATA-1 SMI IMAGES USING STRUCTURAL SIMILARITY MEASURE

J.N.H. Sempio¹, R.K.D. Aranas¹, B.P. Lim¹, B.J. Magallon¹, M.E.A. Tupas², I.A. Ventura¹

¹ STAMINA4Space Program, University of the Philippines, Quezon City, Philippines

² Department of Geodetic Engineering, University of the Philippines, Quezon City, Philippines

Commission IV

KEYWORDS: Remote Sensing, Diwata-1, Georeferencing, Image Transformation, Structural Similarity

ABSTRACT:

This paper aims to provide a qualitative assessment of different image transformation parameters as applied on images taken by the spaceborne multispectral imager (SMI) sensor installed in Diwata-1, the Philippines' first Earth observation microsatellite, with the aim of determining the order of transformation that is sufficient for operationalization purposes. Images of the Palawan area were subjected to different image transformations by manual georeferencing using QGIS 3, and cloud masks generated and applied to remove the effects of clouds. The resulting images were then subjected to structural similarity (SSIM) tests using resampled and cloud masked Landsat 8 images of the same area to generate SSIM indices, which are then used as a quantitative means to assess the best performing transformation. The results of this study point to all transformed images having good SSIM ratings with their Landsat 8 counterparts, indicating that features shown in a Diwata-1 SMI image are structurally similar to the same features in a resampled Landsat 8 data. This implies that for Diwata-1 data processing operationalization purposes, higher order transformations, with the necessary effort to implement them, offer little advantage to lower order counterparts.

1. INTRODUCTION

Whenever a spaceborne imaging system gathers data, it is subjected to several factors leading to geometric distortion such as the Earth's curvature, mismatch in projection, relief displacement, and others. The presence of these geometric distortions cause changes in a satellite image's scale, angular relationships, feature positions, etc., which are needed to be corrected for such an image to become usable for mapping purposes (Dave, Joshi & Srivastava, 2015). As the number of non-systematic distortions present in a satellite image increases, the complexity of the needed corrections via georeferencing also increases. However, there is also the issue of overfitting - the development of an overly complex correction model relative to the amount of data available (Bilger & Manning, 2013) to fit every single detail of the data to the reference. Applying an overfit model will cast doubt in the integrity of the estimates and any new data will cause the model's performance to deteriorate (Bilger & Manning, 2013) - and in the case of satellite images, an overfit georeferencing parameter will have

a detrimental effect on the integrity of pixel values once the data is used in applications such as monitoring the health of vegetation.

This paper explores the implementation of different image georeferencing techniques to correct images coming from Diwata-1, the Philippines' first Earth observation microsatellite, with the aim of determining the most practical transformation algorithm that can be used to minimize cases of oversampling and overfitting of data. In particular, the authors attempt to compare the transformed images with standard data from another satellite.

2. DATA PROCESSING AND ANALYSIS METHODS

2.1. Image transformation

There are several image transformation algorithms that have been developed, but the ones listed in this paper are the techniques present in QGIS 3, being the software used to georeference and transform the Diwata-1 SMI images.

2.1.1. Affine/Helmert: An affine transformation (equivalent to the "Helmert transformation" in QGIS,

which is a subset of the affine transformations) is a type of transformation that preserves points, straight lines and planes as long as the data are preserved in affine space (i.e. changes can be defined by simple translations or displacement vectors) (Fisher, Perkins, Walker & Wolfart, 2003). Thus, affine transformation modifies the data primarily via translation and rotation (Euclidean), scaling (similarity) and uniform planar shearing (affine) (Fisher, Perkins, Walker & Wolfart, 2003), and ultimately performs the least amount of warping on an image. Consequently, affine transformation does not consider distortions due to the Earth's curvature and the terrain of the imaged area and is thus most suitable for images where the effects are minimal, such as images of flat surfaces with small area coverage.

2.1.2. Polynomial: When the effects of rough terrain or of the Earth's curvature are no longer negligible, affine transformation is not going to be suitable to obtain a proper matching of features with the reference basemap. Thus, it is going to be necessary to warp the image in a manner akin to "rubber sheeting", and polynomial transforms are going to be needed.

A polynomial transformation is a non-linear transformation and relates two 2D Cartesian coordinate systems through a translation, a rotation and a *variable scale change* (Knippers, 2009) - this last item is what differentiates polynomial transformations from its affine counterpart. The transformation function can have an infinite number of terms, but QGIS 3 can perform up to the third degree.

2.1.3. Projective: A transformation that maps lines to lines - but not necessarily preserve parallelism - is called a **projective transformation** (Levy, 1995). This type of transformation considers the projection of a feature from one plane of view into another plane of view, similar to how a single object can have different shapes when a camera takes an image from different angles.

2.1.4. Thin Plate Spline: The **thin plate spline** (TPS) algorithm is a georeferencing method based on a spline interpolation - "a piecewise polynomial that maintains continuity and smoothness between adjacent polynomials" (Ha, et al., 2018). Because of its ability to generate a smooth function that

interconnects separate piecewise polynomial functions, spline-based transforms allow a much greater range of image warping options than higher order polynomial functions, resulting in a very high goodness of fit with reference data. There are two downsides on the use of TPS, however: the need to generate a much larger number of GCPs than polynomial functions, and the increased risk of overfitting the data due to the number of interconnected polynomial functions.

2.1.5. Remarks: In Diwata-1 operations, when manual georeferencing is performed in QGIS the minimum number of GCPs needed to generate reliable residual calculations increases with the order of transformation. When GCPs are evenly distributed across the image being worked on, a Helmert transformation needs at least three points, polynomial transformations at least seven, projective transformation also at least seven, and TPS at least twenty. The number of GCPs will increase further depending on residual calculations because some initially located points must be removed and replaced by more properly located points to improve the residual score and thus ensuring the image's goodness of fit when overlaid on a basemap. As a georeferencing task on a single image takes ten to fifteen minutes to complete with the minimum needed transformation, the same task using a higher transformation will take even longer with possible minimal gain in accuracy. This study will therefore consider not only the visual accuracy of the transformed image but also the potential time expended on a georeferencing task, in order to determine the optimal transformation that can be used that will also help save time for operations purposes.

2.2. Cloud Masking

The presence of clouds and cloud shadows in a satellite image reduces its usability for purposes of thematic mapping, but there may still be useful information that can be derived from non-cloud areas. When clouds are present in an image intended for data analysis, a mask is created to minimize, if not remove, the effects of clouds by setting their values to a constant value, normally to zero.

Using a cloud mask also helps in data comparison activities, as the implementation of an image's cloud mask on the item it is going to be compared with

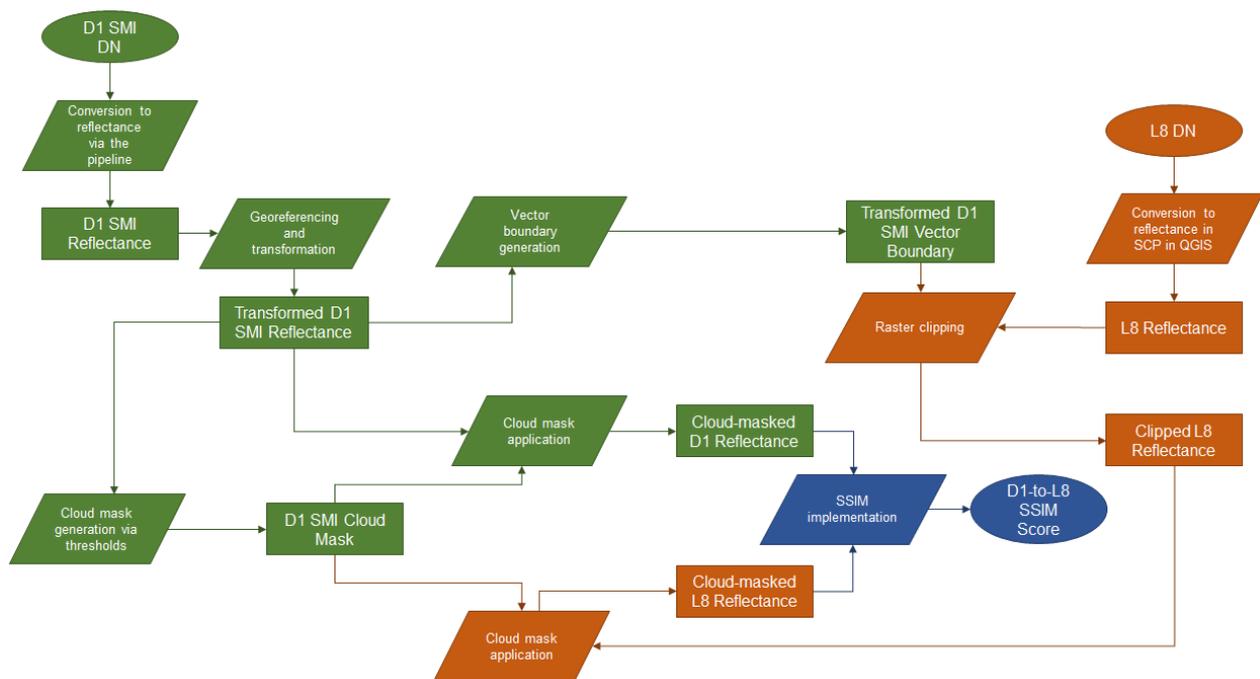


Figure 1. Flowchart of this study's methodology

eliminates the otherwise positive bias that will occur in favor of the otherwise cloud-free image.

Several cloud mask generation techniques were researched to find the most applicable one for Diwata-1 data, including the Otsu process (Ramos et al., 2016) and the use of reflectance-based thresholds via adaptive average brightness (Roomi, Bhargavi & Banu, 2012). For the purpose of this paper, the threshold technique is used.

2.3. Structural Similarity

The structural similarity index (SSIM) is a full reference metric that gives a numerical rating on the perceived similarity of a test image with a reference image (Wang, Bovik, Sheikh & Simoncelli, 2004). In contrast with other quality metrics such as peak signal-to-noise ratio, that give ratings based on the entire scene, SSIM assesses the scene in terms of *structural information*, giving ratings based on the closeness of features to the reference. As SSIM measures the perceptual difference between two images (Wang, Bovik, Sheikh & Simoncelli, 2004), it does *not* determine which of the two images is better. This implies the need for a suitable reference image when assessing the test image via SSIM, and thus this study

made use of Landsat 8 as reference data, given the Landsat Program's long history of operations.

Two SSIM implementations have been tested for this study, with the goal of obtaining the most reasonable SSIM index output. One is an implementation as a function in the Python-based scikit-image processing library ([link](#)), and the other is a standalone Python code made by Oliveira ([link](#)). From the said tests, the Oliveira code was deemed better than the scikit-image function because the resulting SSIM indices for the former fall within the zero to one range of values, whereas the latter was giving values exceeding one.

3. IMPLEMENTATION ON DIWATA-1 DATA

3.1. Preparation of Diwata-1 data

QGIS 3.6 was used to perform image-to-image registration on Diwata-1 images of Sofronio, Palawan with the aid of Landsat 8 data, ensuring that the maximum calculated residual value for a ground control point (GCP) was below one-pixel unit. New images were then generated by transforming the test images via Helmert, Polynomials 1, 2 and 3, Projective and Thin Plate Spline transformations.

After transformation, the images' DN values were converted into their reflectance equivalents. To ensure uniformity when SSIM is implemented at the end, the

SMI data were resampled to 60m spatial resolution. Cloud masks were then generated from the resampled data using the threshold technique as programmed in the Python language. These masks were then in turn applied on the SMI images using QGIS 3.6 and the Python-based Geospatial Data Abstraction Library (GDAL).

Vector boundaries of the georeferenced images were also generated for clipping scenes from the reference Landsat 8 data, discussed in the next section.

Table 1: Number of manually obtained ground control points (GCPs) used for each transformation type of Diwata-1 SMI data, with the guideline that the obtained residual value is below one-pixel unit. This was obtained by reprocessing the GCPs used for the TPS transformation to become usable by the lower order transformations.

Transformation Type	V490	V550	V670	V690
Helmert	78	43	41	53
Polynomial 1	78	63	60	53
Polynomial 2	78	63	60	53
Polynomial 3	78	63	60	53
Projective	85	63	60	52
Thin Plate Spline	85	63	60	52

3.2. Preparation of Landsat 8 data

Landsat 8 images matching the season and central wavelength of the Sofronio data were downloaded via the USGS EarthExplorer website ([link](#)) to be used as reference data for SSIM. To match the reflectance measures in Diwata-1 SMI images, the downloaded data's pixel values were converted from DN to reflectance with the aid of the preprocessing tools made available by the Semi-automatic Classification Plugin module for QGIS 3.6 ([link](#)). The reflectance-converted Landsat 8 images were in turn subjected to clipping using the generated boundaries from georeferenced Diwata-1 SMI data, and then resampled to 60m so that their spatial resolution will match their Diwata-1 SMI equivalents.

Table 2: The value of central wavelengths of Diwata-1 SMI images used in the study, and their Landsat 8 counterparts.

Diwata-1 SMI	Landsat 8
V490	Band 2
V550	Band 3
V670	Band 4
V690	Band 4

Cloud masks were then applied on the extracted Landsat 8 images in QGIS and GDAL to match the cloud presence of their respective Diwata-1 SMI counterparts. And if the obtained Landsat 8 data also has clouds present, a mask is generated from these cloud data and then applied to their corresponding Diwata-1 SMI counterparts (in the case of this study, the obtained Landsat 8 data of the study area were cloud free, and this step was unnecessary).

3.3 Testing via SSIM

Once the respective pairs of Diwata-1 SMI and Landsat 8 images had been prepared, they were subjected into SSIM analysis via the sourced Oliviera code. The resulting SSIM scores for each pair were then tabulated.

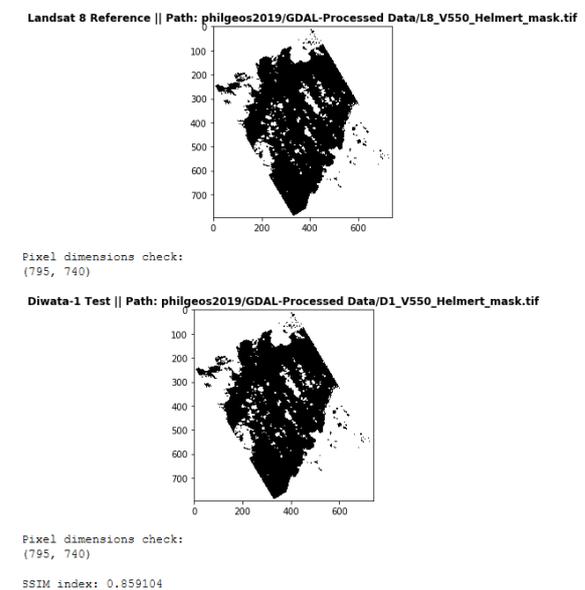


Figure 2. Screenshot of the output of the prepared code for SSIM implementation on Diwata-1 SMI V550 image of Sofronio, Palawan, warped via Helmert transformation

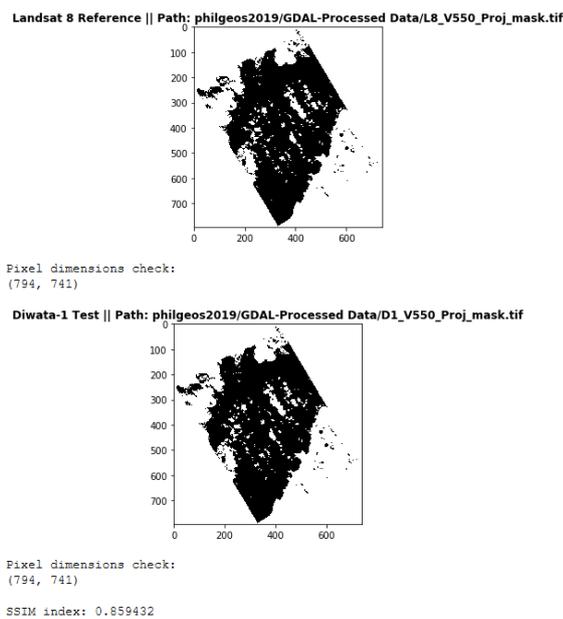


Fig.3. Screenshot of the output of the prepared code for SSIM implementation on Diwata-1 SMI V550 image of Sofronio, Palawan, warped via projective transformation

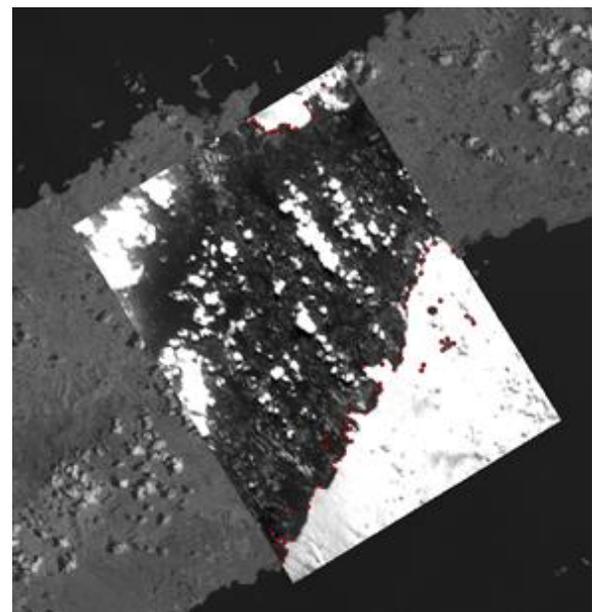


Fig.4. A georeferenced Diwata-1 image of the Sofronio area in Palawan, taken in 06 July 2017 at 12:01 PM (GMT+8), overlaid on a single band Landsat 8 image. Note the very bright water areas.

Table 3: Recorded SSIM indices for each pair of transformed Diwata-1 image and Landsat 8 counterpart.

Transformation Type	V490 (mask = 0.15)	V550 (mask = 0.30)	V670 (mask = 0.30)	V690 (mask = 0.30)
Helmert	0.889791	0.859104	0.885785	0.894472
Polynomial 1	0.811503	0.855832	0.889657	0.893728
Polynomial 2	0.878789	0.859595	0.889803	0.886411
Polynomial 3	0.817725	0.85945	0.887075	0.889536
Projective	0.874502	0.859432	0.890642	0.887871
Thin Plate Spline	0.877647	0.858703	0.886737	0.887964

4. OBSERVATIONS

4.1. Analysis of results

One of the most notable issues regarding the sample Diwata-1 SMI data from Sofronio, Palawan is the presence of glare on water bodies, being very bright areas instead of the expected darker pixels. This, in turn, results in the generated threshold-based cloud mask to accidentally include these water areas as “cloudy”, besides actual clouds over land and water. This situation proved fortunate, as water features were also removed for both Diwata-1 and Landsat 8 data after application of the cloud mask, virtually leaving land features to be the areas assessed by SSIM.

Table 3 presents the generated SSIM indices for the respective pairs of Diwata-1 SMI and Landsat 8 imagery, with its corresponding bar chart in Fig.5.

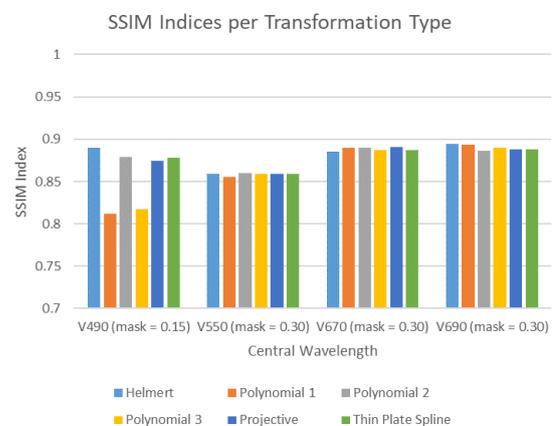


Fig.5. A bar chart representation of the values in Table 3.

Except for the V490 pair, all calculated SSIM indices range between 0.85 to 0.9, denoting that a Diwata-1 SMI image may be structurally similar to that of a resampled Landsat 8 data. In other words, the depiction of features in a Diwata-1 SMI image may be comparable to how a resampled Landsat 8 image presents them. This result deviated from the authors' expectations of observing lower SSIM indices for lower order transformations.

Table 4: Compiled SSIM indices for salt-and-pepper noise from different studies

Study	SSIM Rating
Lee & Lim, 2016 (Lena)	0.297
Lee & Lim, 2016 (Einstein)	0.781
Wang, Wang, Liao & Lim, 2008	0.6085
Lu, Bi & Wang, 2010 (White Noise)	0.7062
Yu & Liu, 2011 (Goldhill, $i=2$)	0.72

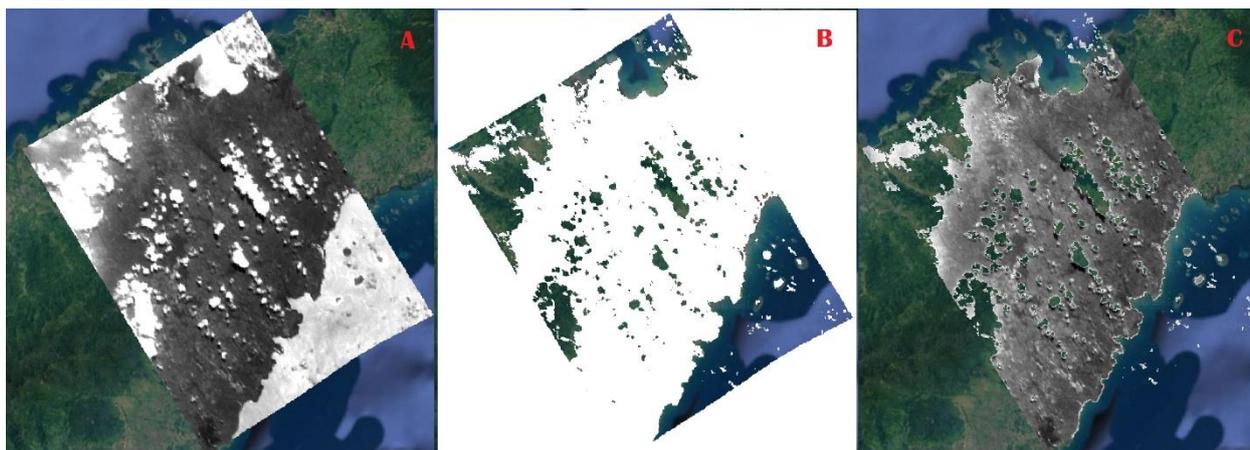


Fig.6. An example of an application of a cloud mask. Item (A) is a processed Diwata-1 image of the Sofronio area, transformed using thin plate spline. Item (B) is its corresponding cloud mask, where the white areas are valued as '1' and the transparent areas (denoting the cloudy areas and the bright water areas) are valued as '0'. Item (C) is the resulting output using Raster Calculator in QGIS 3.6.

To determine if the deviations observed in the bar chart for the V490 pair is significant, a reference SSIM index-based rating system is needed. This study attempted to find researches that present an itemized interpretation of SSIM indices to determine the most acceptable index value but was unsuccessful in this task. The researchers thus resorted to an indirect method by compiling several studies attempting to relate SSIM with subjective human visual perception (Lee & Lim, 2016; Wang, Wang, Liao & Lim, 2008; Lu, Bi & Wang, 2010; Yu & Liu, 2011) and obtaining the indices calculated from images with salt-and-pepper noise, as this item is given low degradation mean opinion scores (DMOS) by human observers (Minamoto & Ohmura, 2014). From these studies, it was inferred that an SSIM rating of 0.8 can be considered as an acceptable similarity measure so that noisy images will fall below it.

With all recorded SSIM ratings in this study being above 0.8, it can be inferred that the resampled Diwata-1 SMI data, regardless of the degree of transformation, are comparable with their resampled Landsat 8 counterparts. For operationalization purposes, this observation is important, as it implies there is minimal advantage to using higher order transformations to generate an output, considering the effort needed to implement them.

4.2. Issues in the Analyses

Perhaps the most pressing issue in the observed results is that the conclusion on georeferencing outputs being structurally similar with the reference data is made on top of an assumed means of interpreting the SSIM scores. Further research is going to be needed to reinforce this assumption.

Another potential issue is the question of the SSIM index being an appropriate measure for the analysis made in this research. To test this angle, the SSIM code was run to compare a raw pair of Diwata-1 SMI and Landsat 8 images *without applying the cloud mask*. The resulting SSIM index was valued at

0.135243, denoting very low similarity between the two images, and is the expected outcome.

To further check on the validity of the SSIM code, the application was run by using a Helmert transformation of a Diwata-1 SMI image as reference and a TPS transformation of *the same image* as test. And to push the testing further, a reference image is compared with itself. The former test scored a high SSIM index at 0.877084, and the latter with a perfect 1.0, both results being expected. And as a final check, a raw Diwata-1 SMI image was tested against the same image with its cloud mask applied, and the resulting SSIM index is low at 0.242294, which is once again expected. The results of the additional tests established the robustness of the SSIM code used in the study and reinforces the validity of the outputs of Table 3.

And another potential issue is that SSIM may be misinterpreted as a measure of accuracy of a georeferenced image. It is important to emphasize that SSIM can only assess the similarity of features presented between a reference and a test image. The indices do not determine, in any way, the correctness of image registration against a base map, and thus it is still up to a quality control checker to ensure that the image is correctly registered when overlaid on a base map. This issue is particularly relevant for the otherwise high ratings obtained by the Helmert-transformed Diwata-1 SMI data.

5. CONCLUSIONS

This study was aimed to determine the most practical georeferencing transformation method that can be used for Diwata-1 SMI images by comparing the different transformation algorithms with an equivalent Landsat 8 image. The structural similarity (SSIM) index was used to generate a quantitative rating for each transformed image with which the abovementioned goal can be achieved.

Because the SSIM indices of the test pairs were above the 0.8 threshold, it can be concluded that the georeferenced Diwata-1 SMI outputs using different transformation parameters are structurally similar to the reference resampled Landsat 8 data. And because it takes longer to manually generate a transformed image with higher order algorithms, based on the SSIM conclusions, it can be declared that warping an image up to the second polynomial will suffice for operationalization of georeferencing activities.

As stated in the previous section, further tests are needed to reinforce the validity of the outputs of the research. This study did, on the other hand, provide a working methodology that allows practical usage of SSIM in determining the comparability of a sample satellite image with an established reference, which in the latter's case is with Landsat 8 imagery. This study therefore recommends testing the methodology with other datasets using different test data (such as from Diwata-2) or using a different reference (such as from Sentinel-2).

The authors of the study would also like to note that to improve the reliability of the methodology, the test data may need further refining before being subjected to SSIM analysis. This can be achieved by applying relief displacement corrections on the test image using a digital elevation model (DEM), and by considering the application of spectral band adjustment to match the central wavelength of the test image with the ones used in the reference image.

ACKNOWLEDGMENTS

This research is made possible by using data generated by the STAMINA4Space Program, which is funded by the Department of Science and Technology (DOST), monitored by DOST-Philippine Council for Industry and Emerging Technology Research and Development (PCIEERD), and implemented through the collaboration between the University of the Philippines – Diliman, the Advanced Science and Technology Institute (ASTI) of the DOST, and Hokkaido University and Tohoku University in Japan. The authors would also like to acknowledge the contributions of Benjamin Jiao and Mark Jayson Felix by providing insights to help address issues encountered in data processing.

REFERENCES

- Bilger, M., & Manning, W.G. (2013). Measuring overfitting in nonlinear models: A new method and an application to health expenditures. *Health Economics*, 24(1), pp. 75 - 85.
- Dave, C. P., Joshi, R., & Srivastava, S.S. (2015). A Survey on Geometric Correction of Satellite Imagery.

International Journal of Computer Applications, 116(12), pp. 24 - 27.

Fisher, R., Perkins, S., Walker, A., & Wolfart, E. Affine Transformation. Retrieved from Hypermedia Image Processing Reference: <http://homepages.inf.ed.ac.uk/rbf/HIPR2/affine.htm>

Ha, P.V., Thanh, N.T.N., Hung, B.Q., Klein, P., Jourdan, A., & Laffly, D. (2018). Assessment of georeferencing methods on MODIS Terra/Aqua and VIIRS NPP satellite images in Vietnam. 10th International Conference on Knowledge and Systems Engineering (10th KSE) (pp. 282 - 287). Ho Chi Minh City: IEEE.

Knippers, R. (2009). Coordinate transformations. Retrieved from the ITC-GIP's Geometric Aspects of Mapping Page: <https://kartoweb.itc.nl/geometrics/Coordinate%20transformations/coordtrans.html>.

Lee, D., & Lim, S. (2016). Improved structural similarity metric for the visible quality measurement of images. Journal of Electronic Imaging, 25(6). Retrieved from the SPIE Digital Library: <https://www.spiedigitallibrary.org/journals/Journal-of-electronic-imaging/volume-25/issue-06/063015/Improved-structural-similarity-metric-for-the-visible-quality-measurement-of/10.1117/1.JEI.25.6.063015.full?SSO=1>

Levy, S. (1995). Projective Transformations. Retrieved from the University of Minnesota Geometry Center Home Page Archive: <http://www.geom.uiuc.edu/docs/reference/CRC-formulas/node16.html>. Content excerpted from the 30th Edition of the *CRC Standard Mathematical Tables and Formulas* by the CRC Press.

Lu, D., Bi, D., & Wang, Y. (2010). Image Quality Assessment Based on DCT and Structural Similarity. 6th International Conference on Wireless Communications Networking and Mobile Computing (WiCOM). Chengdu: IEEE.

Minamoto, T., & Ohmura, H. (2014). Indices for image quality degradation evaluation based on wavelet transforms. Proceedings of the 2014

International Conference on Wavelet Analysis and Pattern Recognition (pp. 146 - 152). Lanzhou: IEEE

Ramos, M. G., Jiao, B. J., Aranas, R. K., Magallon, B. J., Amado, J., Tupas, M. E., & Tamondong, A. (2016). Automated calculation of cloud cover from RGB composite of Landsat 8 and Diwata-1 satellite imagery. 37th Asian Conference on Remote Sensing (pp. 1 - 9). Colombo: ACRS.

Roomi, M. M., Bhargavi, R., & Banu, T. M. H. R. (2012). Automatic identification of cloud cover regions using SURF. International Journal of Computer Science, Engineering and Information Technology, 2(2), pp. 159 - 175.

Wang, B., Wang, Z., Liao, Y., & Lin, X. (2008). HVS-based Structural Similarity for Image Quality Assessment. 2008 9th International Conference on Signal Processing. Beijing: IEEE.

Wang, Z., Bovik, A.C., Sheikh, H.R., & Simoncelli, E.P. (2004). Image quality assessment: from error visibility to structural similarity. IEEE Transactions on Image Processing, 13(4), pp. 600 - 612.

Yu, H. & Liu, X. (2011). Structure similarity image quality assessment based on visual perception. Proceedings of 2011 International Conference on Electronic & Mechanical Engineering and Information Technology. Harbin: IEEE.