

RADIOMETRIC CALIBRATION OF LANDSAT-8 OLI AND SENTINEL-2 MSI IMAGES FOR WATER QUALITY MODELLING OF LAGUNA LAKE, PHILIPPINES

E. V. Gubatanga Jr^{1,*}, A. C. Blanco¹, C. H. Lin², B. Y. Lin²

¹Training Center for Applied Geodesy and Photogrammetry, University of the Philippines, Diliman

²Department of Geomatics, National Cheng Kung University, Tainan City, Taiwan
evgubatanga@gmail.com, acblanco@up.edu.ph, linhung@mail.ncku.edu.tw, pikadp2004@hotmail.com

Commission IV

KEYWORDS: canonical correlation analysis, pseudo-invariant features, multivariate alteration detection

ABSTRACT:

Regular monitoring of water quality in Laguna Lake is important for it supports aquaculture and provides water supply for Metro Manila. Remote sensing makes it possible to monitor the spectral conditions of the lake on a regular time interval and with complete coverage except for the areas with cloud and shadow cover. Along with in-situ water quality measurements, bio-optical models can be developed to determine the relationship between spectral and bio-optical properties of the lake water and consequently enables the estimation of water quality through remote sensing. However, radiometric calibration is needed to minimize the effects of the changing atmospheric conditions over time and to account for the difference in sensors (e.g., Landsat-8 OLI, Sentinel-2 MSI) used for water quality assessment. Canonical correlation analysis is used to detect pseudo-invariant features (PIFs), which are ground objects that do not dramatically vary in spectral properties over time. Road surface and other large man-made infrastructures are the commonly detected PIFs. These PIFs are used to compute for the parameters used to normalize reflectance values of remotely-sensed images obtained on different dates and using different sensors. The normalization resulted to a reduction of difference in reflectance values between the reference image and the adjusted image, though not marginal. This is due to the use of a linear equation to adjust the image, which limits the ability of the reflectance values of the image to fit to the values of the reference image.

1. INTRODUCTION

Remote sensing is a technology that acquires earth surface data using a sensor attached to a platform such as an aircraft or a satellite. Depending on remote sensing system's temporal resolution, the sensor revisits the same portion of the earth surface after a certain period of time. Atmospheric conditions vary on every satellite revisit. These variations affect the change measurements on the earth surface (Denaro, et al, 2018). Therefore, it is important that these variations be minimized to obtain a more accurate assessment of the earth cover change. Radiometric calibration aims to solve this problem. It works by searching for pseudo-invariant features (PIFs) which are ground objects that tend to retain their spectral properties over time. Examples of these are roads, roofs and large rocks. A direct subtraction of the reference image from the target image can be used to find PIFs. However, PIFs that are obtained using this method are not optimal because the two images can be further differentiated. The optimal PIFs are detected by maximizing the variance between two images consisting of a reference image and the image to be adjusted. This can be done through canonical correlation analysis. It determines the coefficients that can be multiplied to the arrays representing the images and will produce linear combinations with maximum variance. Better PIFs can be detected using this method and will provide better adjustment parameters (Canty, et. al, 2004).

The University of the Philippines Multi-platform and Cross-sensor Water Quality Monitoring (UP MCWQM) Project aims to monitor the water quality across the entire Laguna Lake through the use of remote sensing and water quality modelling. The project proposes an alternative to costly and labour-intensive in-situ sampling for water quality measurements that only covers several points in the lake. Through remote sensing, water quality measurements across the entire lake can be derived which gives

way to a much detailed spatial and temporal analysis of water quality. Reflectance values provided by the satellite images can be used as independent variables for the computation of water quality parameters such as turbidity and chlorophyll-a concentration (Barrett & Frazier, 2016). Various water quality models such as the BOMBER bio-optical model (Giardino, et. al, 2012) and WASI-2D software (Gege, 2014) are used for this task. In-situ measurements are used for the calibration and validation of these models. The project also looks into the potential of using remotely-sensed images from an unmanned aerial system (UAS) for water quality modelling.

An important step prior to water quality modelling is the pre-processing of the satellite images. It involves atmospheric correction, cloud and shadow masking and radiometric calibration. The project uses multivariate alteration detection (MAD) method to detect PIFs for radiometric calibration. MAD uses the aforementioned canonical correlation analysis to maximize the difference between 2 images to find the optimal PIFs. MAD is sensitive to the presence of clouds because MAD can misclassify clouds as PIFs due to their consistent appearance. If a certain location captured by the satellite image that is covered with clouds for both its target and reference image, there is a high probability that it will be mistaken as a PIF (Lin, et. al, 2017). Due to this problem, novel approaches such as the iteratively reweighted MAD (IR-MAD) are developed (Nielsen, 2007). However, since cloud and shadow masking precedes radiometric calibration in the project, a simple MAD approach could be sufficient. This research investigates the effectivity of using MAD for radiometric calibration if the images that are to be calibrated are cloud and shadow-masked beforehand.

2. METHODOLOGY

2.1 Image Pre-processing

There are 3 satellite platforms used in the study: Sentinel-2, Landsat-7 and Landsat-8. Images undergo atmospheric correction according to which sensor is used. Sentinel-2 uses sen2cor while Landsat-7 and Landsat-8 both use FLAASH. After atmospheric correction, the images will undergo cloud and shadow masking using Tmask algorithm (Zhu, Z. & Woodcock, C., 2014). Pixels that are detected to have cloud and shadow cover are then replaced with pixel value of zero across all bands. Along with the masking process is the cropping of the images using the bounding box of the Laguna Lake's watershed.

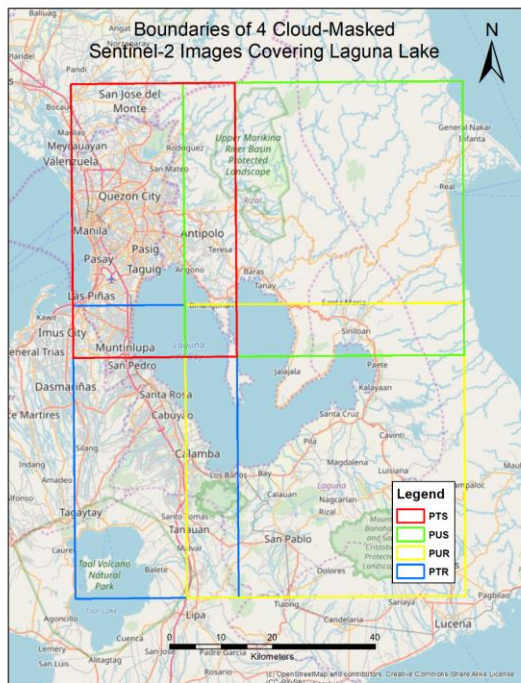


Figure 1 boundaries of 4 cloud-masked Sentinel-2 images covering Laguna Lake

One issue regarding Sentinel-2 is that it takes 4 scenes to complete the coverage of Laguna Lake while both Landsat-7 and Landsat-8 only takes 1. Therefore, it will be necessary for the 4 scenes of Sentinel-2 to be mosaicked into a single image before calibration. Sentinel-2 follows a naming convention to name the tiles according to their position. For Laguna Lake, the images that cover it have index names PTR, PTS, PUR and PUS. The PUR and PUS images are acquired by the Sentinel-2 satellite consecutively then returns after a short period of time to acquire PTR and PTS images. Since the time difference between the acquisition of PUR and PUS images is very small, the difference between raw pixel values of the 2 images within their overlap will be very small. Therefore, no radiometric calibration will be needed between PUR and PUS image pair. The PTS and PTR image pair also has the same condition. However, the time difference between the acquisition of PTS and PUS images are significant that it requires radiometric calibration. It is the same case for PTR and PUR images. The process of radiometric calibration will be discussed in the succeeding chapter. But first, PTS and PTR will be mosaicked by feathering. Feathering will also be used for mosaicking PUS to PUR. Then, radiometric calibration will be done between the PUS – PUR image pair and the PTS – PTR image pair.

During the feathering process, the DN values in the overlap are computed by multiplying the DN values from PTS and PTR to their respective weights and then adding the 2 products together. The weight is based on the position of the pixel in the overlap. PTS pixels have higher weights over PTR pixels in locations closer to PTS, which are basically on the top portion of the overlap. On the other hand, PTR pixels will have higher weights in the lower portion of the overlap. The same scheme is done for the PUS – PUR pair.

Two image mosaics are then produced, one on the left side and the other on the right. As discussed earlier, these pairs cannot be directly mosaicked since there is significant time difference on the time of acquisition between these images and will require radiometric calibration prior to mosaicking. The right side, the PUS – PUR pair, will serve as the reference image while the left side, the PTS – PTR pair, will be the one to be adjusted. The reason for this is because the PUS – PUR pair cover the greater portion of the Laguna Lake and choosing the PUS – PUR pair as the reference image will lessen the number of pixels that need adjustment.

Radiometric calibration involves detection of PIFs. After the detection of PIFs, the adjustment parameters are derived and the PTS – PTR pair is adjusted. Then, the two pair are adjusted. But in this case, instead of using feathering, the PUS – PUR pair, which is the reference image, is given priority in the overlap. If a PUS – PUR pixel is present (because it is not covered by cloud or shadow), regardless if there is also a PTS – PTR pixel present in the same position, the final mosaic will adopt the values of the pixel from the reference image. The only time that the final mosaic will adopt a pixel from the adjusted PTS – PTR pair is when there is no PUS – PUR pixel available on a specific location.

2.2 Detection of Pseudo-Invariant Features

The difference between two images can be directly obtained by simply subtracting the target image from the reference image. The pixels with small difference can be considered as PIFs and their difference values can be the basis for the image adjustment. However, this is not the optimal solution since the difference between two images can further be increased using canonical correlation analysis (Nielsen, et al., 1998). Maximizing the difference between the two images results in the detection of the optimal PIFs for adjustment.

An image is converted first to a numpy array. This numpy array is 3-dimensional wherein the first dimension is the number of bands while the other two dimensions represent the height and the width of the images. The numpy array is then reduced to 2 dimensions by flattening each 2-d numpy band into a single row of pixels. The new dimensions of the new numpy array will be the number of bands by the number of pixels. Before the detection of pseudo-invariant features, the cloud and shadow masks are first detected to be excluded during the computation.

The numpy arrays will then be used to compute for the variance and covariance of the images. The product of the reference and the target image is the covariance. The reference image multiplied to itself is its variance. The variance of the target image is also obtained. The square of the variance is then multiplied to the inverses of variance of the reference and target image. The resulting matrices are shown in the equations below.

$$\sum_{XX}^{-1} \sum_{XY} \sum_{YY}^{-1} \sum_{YX} = \rho^2 a \quad (1)$$

$$\sum_{YY}^{-1} \sum_{YX} \sum_{XX}^{-1} \sum_{XX} = \rho^2 b \quad (2)$$

\sum_{XX} and \sum_{XY} are the variances of the reference and the target images, respectively. \sum_{XX} and \sum_{XY} on the other hand represent the covariance between the 2 images. The resulting matrix will undergo Eigenvalue decomposition in order to obtain coefficients a and b and the correlation ρ . The number of eigenvalue – eigenvector pair will be equal to the number of bands. In the case of Sentinel-2, there will be 9 eigenvalue – eigenvector pairs. Each eigenvalue represents a square of correlation between the reference and target image. Each of the reference and the target image will have its own eigenvector. The lowest eigenvalue corresponds to the eigenvectors that produce the highest variance between the reference and the target image. Each element in the eigenvector will be multiplied to each band of the image. The resulting product is called the linear combination and is represented by variables U and V for the reference and target images, respectively. The linear combination of the target image will be subtracted to the linear combination of the reference image as shown on the equation below where i refers to the band number.

$$MAD_i = U_i - V_i = a_i^T X_i - b_i^T Y_i \quad (3)$$

The sum of the resulting differences is then normalized using the standard deviation. PIFs are determined using the normalized sum. A chi-squared test is applied wherein the normalized sum is compared against the threshold. Pixels with normalized sum above the threshold are considered as PIFs (Wang & Lin, 2016).

$$\sum_{i=1}^p \left(\frac{MAD_i}{\sigma_{MAD_i}} \right)^2 > \chi_{\alpha,df}^2 \quad (4)$$

PIFs are then divided into 2 sets: training and test sets. The training set is used to derive the adjustment parameters. The DN values of reference and target image within the training PIFs are gathered across all bands. The standard deviation of these DN values are calculated. The ratio between the standard deviation derived from the target image and the standard deviation from the reference image will serve as the slope that will be used for the adjustment. The intercept is also determined using the equation.

It is preferred that the reference image will be the same for Landat-7, Landsat-8 and Sentinel-2. Having a separate reference image for each platform will prevent cross-sensor image analysis. In the adjustment scheme, the reference image will be from Sentinel-2. It is because it has a higher number of bands compared to Landsat-7 and Landsat-8.

Band	Sentinel-2 (reference)	Landsat-8	Landsat-7
Coastal Aerosol		B1	
Blue	B2	B2	B1
Green	B3	B3	B2
Red	B4	B4	B3
NIR	B8A	B5	B4
SWIR 1	B11	B6	B5
SWIR 2	B12	B7	B7
VRE	B5, B6, B7		

Table 1. Band matching among Sentinel-2, Landsat-8 and Landsat-7 images for radiometric calibration

Table 1 shows the scheme used for the matching of the bands for the cross-sensor calibration. The bands listed in the table are based on the bands that are retained after atmospheric correction and cloud and shadow masking. The coastal aerosol band of

Landsat-8 has no corresponding reference image from Sentinel-2. Therefore, a separate image from the coastal aerosol band of Landsat-8 is used as a reference for other coastal aerosol band of other Landsat-8 images.

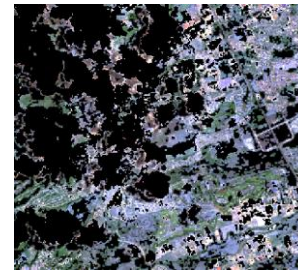


Figure 2 Raw reference image (top); reference image filled with unadjusted image at cloud mask (bottom left); reference image filled with adjusted image at cloud mask (bottom right)

2.3 Image Adjustment

The DN values from both reference and target images of the training PIFs are obtained for the computation of the adjustment parameters. The adjustment parameter computation used is based on the study of (Du, et al, 2002).

$$\alpha_j = \sigma_{Q_{ref}} / \sigma_{Q_j} \quad (5)$$

$$\beta_j = \bar{Q}_{ref} - \alpha \cdot \bar{Q}_j \quad (6)$$

The first adjustment parameter α is measured by dividing the standard deviation of the DN values of the PIFs from the reference values by the standard deviation of the DN values of the PIFs from the target image. Each α is unique per band and j denotes the band number in the equation. It can be interpreted that α is the slope of the standard deviation between the PIFs of the reference and the target image. The second adjustment parameter β is computed using equation 6. \bar{Q}_{ref} is the mean of the DN values of PIFs from the reference image while \bar{Q}_j is the mean of the DN values of the PIFs from the target image. It can be interpreted that β is the intercept. The usage of the adjustment parameters are shown on the equation below.

$$Q_{adj} = \alpha \cdot Q_j + \beta \quad (7)$$

Therefore, the adjustment method proposed by (Du, et al, 2002) is based on a simple linear equation. The training PIFs that are detected can reach up to several hundred-thousand for every pair of images with minimal cloud cover. All of their DN values are used to derive a linear equation. The next chapter discusses if a linear equation is enough to represent the number of PIFs.

In Figure 2, the reference image is shown with cloud and shadow mask. To test the visual improvement regarding the adjustment of the target images, a comparison is made between the unadjusted and adjusted images filling the cloud mask of the reference image. The unadjusted image is darker than the reference image and mosaicking them together may not be as

seamless compared to mosaicking adjusted image with the reference image as seen in Figure 2. It is preferable to do visual comparison in urban areas where more PIFs are present. To quantify the improvement caused by the adjustment, t-test is used.

2.4 Validation

To validate the effectiveness of the radiometric calibration, t-test is used. In a t-test, the level of significance of the difference between the observed measurement and the reference is measured. A low t-test score infers that the difference between the adjusted image and the reference image is statistically insignificant. Statistical insignificance can be interpreted as the difference between the adjusted and the reference image is lower than the magnitude of the random noise in the image. To do this, PIFs are divided into training and test sets. 70% of the PIFs are selected randomly to be part of the training set which will determine the adjustment parameters. The remaining 30% are used to check the effectiveness of the adjustment using a t-test.

Date 2015	t-test scores for unadjusted images			t-test scores for adjusted images		
	blue	green	red	blue	green	red
Jan 25	16.368	-11.346	-5.672	-0.700	-0.307	-1.058
Feb 10	-4.389	-23.223	-11.831	-0.074	1.279	0.763
Feb 26	-	-39.641	-33.776	-0.785	-0.989	-0.729
Mar 14	-9.687	2.695	-0.824	0.169	2.406	-0.411
Mar 30	-	-	-80.418	0.103	0.569	0.827

Table 2. The computed t-test scores of sample unadjusted and adjusted images from Landsat-8.

The training and test sets are selected at random and rerunning the radiometric calibration algorithm will provide the same population of PIFs but slightly different adjustment parameters and t-test scores. There are several cases wherein the adjustment receives a failing t-test score but readjustment causes a new selection of training PIFs and a passing t-test score.

Table 2 shows the t-test scores of sample unadjusted and adjusted Landsat-8 images in the blue, green and red bands. For Landsat-8 which has 7 bands after atmospheric-correction and cloud masking, the t-test score must be greater than -2.365 but less than 2.365 to conclude that the image is statistically the same as the reference. The highlighted cells in Table 2 indicate that a certain band at a certain date did not pass the criteria. It can be seen that none but one band on the sample unadjusted images passed the t-test. Also, the magnitude of several t-test scores of unadjusted images are too large compared to the magnitude of the threshold which is 2.365. Conversely, all of the t-test scores of the sample adjusted image passed except for one. In cases wherein at least one band of an adjusted image has a failed t-test score, the algorithm is re-run until an effective combination of training and test sets of PIFs provides passing t-test scores across all band.

3. CONCLUSION

Basic MAD method is proven to be sufficient to radiometrically calibrate atmospherically-corrected and cloud and shadow-masked Sentinel-2, Landsat-7 and Landsat-8 images. The t-test proves that MAD is able to adjust the target images to have statistically insignificant difference from the target image.

4. REFERENCES

- Barrett, D. C., & Frazier, A. E. (2016). Automated Method for Monitoring Water Quality Using Landsat Imagery. *Water*, 8(6), 257
- Canty, M. J., Nielsen, A. A. & Schimdt, M. (2004). Automatic radiometric normalization of multispectral satellite imagery. *Remote Sensing of Environment*, 91 (3 – 4), 441 – 451
- Denaro L. G., Lin B. Y., Syariz M. A., Jaelani L. M., Lin C. H. (2018) Pseudo-Invariant Feature Selection for Crosssensor Optical Satellite Images. *Remote Sensing & GIS* 7: 239. doi: 10.4172/2469-4134.1000239
- Du, Y., Teillet, P. & Cihlar, J. (2002). Radiometric normalization of multitemporal high-resolution satellite images with quality control for land cover change detection. *Remote Sensing of Environment* 82 (2002) 123 – 134
- Gege, P. (2014). WASI-2D: A software tool for regionally optimized analysis of imaging spectrometer data from deep and shallow waters. *Computers & Geosciences*. 62. 208–215. 10.1016/j.cageo.2013.07.022.
- Giardino, C., Candiani, G., Bresciani, M., Lee, Z., Gagliano, S. & Pepe, M. (2012). BOMBER: A tool for estimating water quality and bottom properties from remote sensing images. *Computers & Geosciences*. 45. 313–318. 10.1016/j.cageo.2011.11.022.
- Lin, B. Y., Wang, Z. J., & Lin, C-H. (2017). Radiometric normalization of multitemporal optical satellite images using iteratively-reweighted multivariate alteration detection. Paper presented at 38th Asian Conference on Remote Sensing - Space Applications: Touching Human Lives, ACRS 2017, New Delhi, India
- Nielsen, A. A. (2007). The Regularized Iteratively Reweighted MAD Method for Change Detection in Multi- and Hyperspectral Data. *IEEE Transactions on Image Processing*, Vol 16, No. 2, February 2007
- Nielsen, A. A., Conradsen, K. & Simpson, J. J. (1998). Multivariate alteration detection (MAD) and MAF postprocessing in multispectral bitemporal image data: New approaches to change detection studies. *Remote Sensing of Environment*, 64(1), 1-19
- Wang, Z., & Lin, C. H. (2016). Pseudo Invariant Features Selection for Optical Satellite Images Using Multitemporal and Multivariate Alteration Detection. *Research NCKU*, 2017
- Zhu, Zhe & Woodcock, Curtis. (2014). Automated cloud, cloud shadow, and snow detection in multitemporal Landsat data: An algorithm designed specifically for monitoring land cover change. *Remote Sensing of Environment*. 152. 217–234. 10.1016/j.rse.2014.06.012.