

CHLOROPHYLL-A CONCENTRATION RETRIEVAL USING CONVOLUTIONAL NEURAL NETWORKS IN LAGUNA LAKE, PHILIPPINES

Muhammad Aldila Syariz (1), Chao-Hung Lin (1), Manh Van Nguyen (1)(4),
Lalu Muhamad Jaelani (2), Ariel C. Blanco (3)

¹ Department of Geomatics, National Cheng Kung University, Taiwan

² Department of Geomatics Engineering, Institut Teknologi Sepuluh Nopember, Indonesia

³ Department of Geodetic Engineering, University of the Philippines, Philippines

⁴ Institute of Geography, Vietnam Academy of Science and Technology, Vietnam

Email: aldilasyariz@gmail.com; linhung@mail.ncku.edu.tw; manh.ig239@gmail.com;
lmjaelani@geodesy.its.ac.id; acblanco@upd.edu.ph

KEY WORDS: chlorophyll-a, estimation, convolutional neural network, Sentinel-3, Laguna Lake.

ABSTRACT: Two existing chlorophyll-a (chl-a) concentration retrieval procedures, which are analytical and empirical, are hindered by the complexity in radiative transfer equation (RTE) and in statistical analyses, respectively. For instance, the analytical model requires basic theory for understanding light propagation in the waters and underlying geophysical parameters which are difficult to obtain in most of the time, making the RTE is inaccurate; while the empirical model selects one best single-, dual-, or even triple- spectral band combination from many possible combination numbers to be regressed to chl-a concentration. Another promising model in this direction is the use of artificial neural networks (ANN). Mostly, a pixel-to-pixel with one-layer ANN model is used; where in fact that the satellite instrumental errors and man-made objects in water bodies might affect the retrieval and should be taken into account. In this study, the mask-based neural structure, called convolutional neural networks (CNN) model containing both the target and neighborhood pixels, is proposed to reduce the influence of the aforementioned premises. The proposed model is an end-to-end multiple-layer model which integrates band expansion, feature extraction, and chl-a estimation into the structure, leading to an optimal chl-a concentration retrieval. In addition to that, a two-stage training is also proposed to solve the problem of insufficient in-situ samples which happens in most of the time. In the first stage, the proposed model is trained by using the chl-a concentration derived from the water product, provided by satellite agency, and is refined with the in-situ samples in the second stage. Eight Sentinel-3 images from different acquisition time and coincide in-situ measurements over Laguna Lake waters of Philippines were utilized to conduct the model training and testing. Based on quantitative accuracy assessment, the proposed method outperformed the existing dual- and triple-bands combinations in chl-a concentration retrieval.

1. INTRODUCTION

Eutrophication refers to the degradation of water quality due to high increasing of phytoplankton biomass in the watershed. It occurs as the aquaculture pond managers intentionally adding fertilizers to reach their goals, that are to enhance primary productivity and to increase the density of important fish (Boyd and Tucker, 1998). Consequently, several aspects have been damaged recently, such as fish kills, human health, and even economic stability. Europe and USA suffered a huge economic losses caused by eutrophication amounting to approximately \$1 billion per year and \$100 million per year, respectively (Hoagland et al., 2012; Shumwey, 1990).

Thus, long-term monitoring and real-time measurements of water quality play a critical role and

recently become worldwide growing concern regarding to the quality of water resources available for multiple resources (Barzegar et al., 2018). Collecting reliable water quality data is an important aspect of water protection and further can be used to better understand the type and severity of water quality impairments. Besides, the water quality monitoring is able to help decision-makers in setting achievable target for water quality improvement.

Conventional approach, that is collecting and analyzing water samples in the professional laboratory, is time-consuming and demanding high cost. For instance, the approach can only represent in-water constituents, such as chlorophyll-a, colored dissolved organic matter (CDOM), etc., at only sampling stations. For the non-station water bodies, which are necessary for long-term monitoring and management, are impossible to obtain (Hajigholizadeh, 2016).

Since 1970's, some researchers have proposed the use of optical remote sensing sensors attached on a satellite to overcome the challenges in conventional approach. This approach soon became a powerful tool that enabled researchers to monitor the in-water constituents in large-scale waterbodies without coming to the sites. A high number of studies has proposed in-water constituent retrieval by utilizing the satellite-derived products such as reflectance and radiance. Two primary retrieval methods are identified, that are empirical and semi-analytical, however, they are suffering complexity problem in statistical and radiative transfer analyses. Another potential method in this direction utilizes neural networks (NN) with multilayer perceptron. This approach is a powerful continuous model to fit multivariate, complex, and non-linear data. (Buckton et al., 1999), (Ioannou et al., 2011), (Ioannou et al., 2013) revealed the NN-based algorithms are known to be able to perform estimation of in-water constituents accurately, both in case 1 and case 2 waters.

Generally, the empirical, semi-analytical, and NN-based method are in common using pixel-to-pixel processing to retrieve in-water constituent concentration with no consideration of neighboring pixels. In fact, satellite images contain both surface reflectance and corresponding errors, including instrumental error. (González Vilas et al., 2011) used a median of 9-pixels mask around the pixel containing the exact geographical location of the sampling point to attenuate the previously mentioned error. This mask could also describe the pixels whether they are closer or further to the boundary of land and water bodies. The in-water constituent concentration near the shoreline has found to be much more than those which is more far away due to the existence of the farming areas and the water inlets surrounding the water body. Thus, the use of the mask is important in order to reduce the effect of instrumental error as well as to define the characteristics of water bodies.

To overcome the problem of the statistical complexity and to attenuate the effect of instrumental error which arise in those previously-mentioned in-water constituent concentration retrieval methods, a deep learning technique with convolutional neural networks (CNN) architecture is adopted and proposed. The networks contain three important stages, that are band expansion, feature extraction, and chlorophyll-a concentration estimation. The three stages have different purposes, that are to cover bands combinations complexity, to describe water bodies characteristics, and to retrieve chlorophyll-a concentration, respectively.

2. STUDY AREA AND DATASETS

Laguna Lake is one of the largest lake in Asia which is located in the southwest of capital city of the Philippines, Manila, with total area of 3,820 km² (Figure 1). The lake features four distinct bays, that are, the West Bay, Central Bay, East Bay, and South Bay (Laguna Lake Development

Authority, 2013). The average depth is 2.8 m and the length of the shoreline is 220 km. More than 100 rivers flow into Laguna Lake with only one output river, that is Pasig River, which is flowing waters to Manila Bay (Santos-Borja and Nepomuceno, 2006).

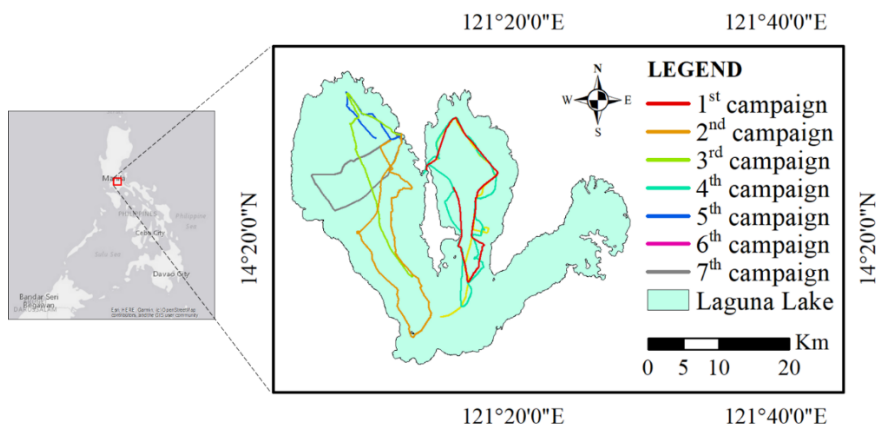


Figure 1. Laguna Lake with tracking line of the seven campaigns

2.1. In-situ Chl-a Concentration Samples

Field data collection campaigns were conducted in seven different days starting from November 2018 to April 2019, covering wet season from June to November and dry season of other months (Maruyama, 2017). The Laguna Lake is divided into three regions, west, center, and east, covering west bay, central bay, and south & east bays, respectively, since it is impossible to cover the whole water bodies in only one campaign. An along-track chl-a data logger was installed on boat to measure chl-a concentration in each seconds. High number of chl-a concentration was recorded; however, several samples link to a pixel in Sentinel-3 imagery because of the image spatial resolution. Thus, an outlier removal and data aggregation were performed in the data pre-processing; and Table 1 lists and summarizes the field campaigns in this study.

Table 1. Summary of field campaigns

Campaign #	Date	Season	Region	# of samples	Chl-a conc. ($\mu\text{g/L}$)
1 st	November 6, 2018	Wet	Center	69	6.8270 ± 0.8369
2 nd	November 14, 2018	Wet	East	40	9.3660 ± 0.8668
3 rd	January 11, 2019	Dry	East	35	11.3386 ± 0.5924
4 th	March 29, 2019	Dry	Center	74	7.9063 ± 0.1646
5 th	April 6, 2019	Dry	East	48	8.4826 ± 1.2292
6 th	April 26, 2019	Dry	Center	22	7.2539 ± 0.3229
7 th	April 30, 2019	Dry	East	98	9.5975 ± 0.8220
All samples					8.6391 ± 1.5383

2.2. Sentinel-3 Imagery

Seven Sentinel-3 water full-resolution (WFR) level-2 images over the study area with the same acquisition dates with the field data collection campaigns were utilized. The images contain not only 16 atmospherically-corrected spectral bands, that is, Band 1-12, 16-17, and 20-21, but also two chl-a concentration channels built by inverse radiative transfer model – neural network (IRTM-NN) and OC4Me which based on the research and algorithm of (Bricaud et al., 1998) and (Morel and Maritorena, 2001). The input to the proposed CNN model is a patch of the size 7x7. Thus, there are 3058, 1712, 1008, 4715, 5681, 1809, and 2582 patches linking to the cloud-free and shadow-free samples in the seven campaign, respectively. The patches containing non-water pixels are excluded from the dataset.

3. METHODOLOGY

3.1. Proposed CNN Model

A convolutional network design, namely CNN model, is chosen and proposed in this study, which is containing three stages: *band expansion*, *feature extraction*, and *chl-a estimation*; which have different purposes which mainly focusing on covering some problems occurred in the remote-sensing approach for chl-a retrieval such as empirical-, analytical-, and ANN-based models. The input to the proposed CNN model is 16-band Sentinel-3 sub-image, excluding spectral band number 13-15 and 18-19 as explained in Section 2.2. The spatial size of the patch image is 7x7 which center and its surrounding pixels belong to water bodies. As for output, the model produces the estimation of chl-a concentration in the center pixel only. This indicates that the model is an image-to-pixel modelling. In order to consider both spectral and spatial information, 3-D filter mask design is utilized instead of 2-D. Figure 3 illustrates the network structure of the proposed CNN model for chl-a concentration retrieval. As a result, there are 4753 unknown parameters within the proposed CNN model and their breakdowns are explained in the next paragraphs.

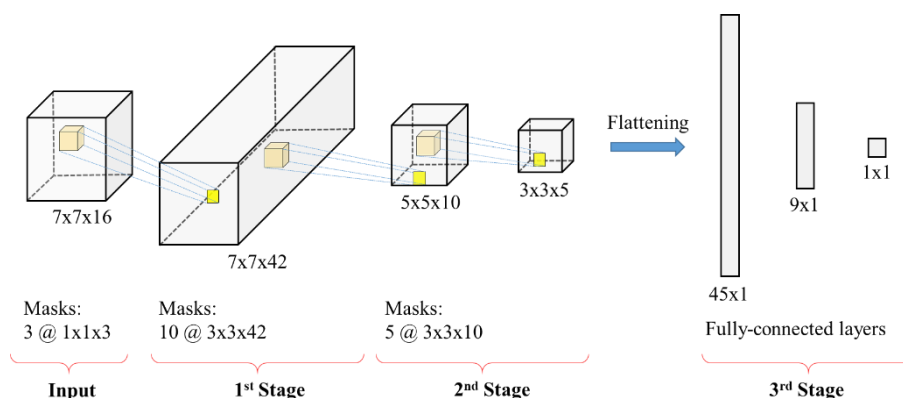


Figure 3. Network structure of proposed model

Band expansion stage. This stage aims to defeat the problem arise in the empirical- and analytical-based models where the complexity was occurred in both approaches. It attempts to deepen the depth of spectral information by combining the spectral bands by means of a set of three 1x1x3-sized filter masks. Each filter mask, combined with bias, utilizes rectified linear unit (ReLU) function for activation function and then followed by batch normalization to further maintain the mean and standard deviation of the activation close to 0 and 1, respectively. There are 24 unknown parameters within the first stage; 12 is covering the number of weights and biases in the filter masks while the other makes up the values of mean, standard deviation, shift, and scaling in batch normalization. Moreover, as no padding is necessary, a single filter mask yields 14 feature images with similar spatial size to the input; meaning that, with three filter masks, 42 7x7-sized feature images are obtained.

Feature extraction stage. This stage contains two convolutional layer in which one layer attempts to reduce the effect from the instrumental errors while the other is for in-water object detection. In the first layer, ten filter masks are utilized with size of 3x3x42, producing 10 feature images after conducting activation function and batch normalization with similar adjustment to the first stage. The spatial size of the feature images in the first layer is 5x5. Moreover, the second layer utilizes five filter masks where each filter mask is sized 3x3x10. The adjustment of activation function and batch normalization is similar to previous layer. The products within this layer are five feature images of the spatial size 3x3. In total, there are 4,305 unknown parameters within this stage; covering 3,830 unknown parameters in the first layer and 475 unknown parameters in

the second layer.

Chl-a estimation stage. Different to the previous two stages, this stage contains flattening layer to reshape 3-D image to a 1-D vector information. As the feature images from previous layer in the second stage is sized $3 \times 3 \times 5$, the vector contains 45 different information. The vector information is connected to a hidden layer with 9 neurons and then the information in the hidden layer is used to estimate chl-a concentration at the center pixel of the patch in output layer. A sigmoid function is used for activation function in both hidden and output layers. In total, there are 424 unknown parameters covering the values of weights and biases connecting the vector information to hidden layer and hidden layer to output layer.

3.2. Proposed Two-stage Training

In order to cover the problem of insufficient samples in model training, a two-stage training is proposed in this study. In the first stage, the proposed CNN model is pre-trained by using chl-a concentration derived from existing retrievals. There are two candidates of retrieval, that are IRTM-NN and OC4Me which are included in Sentinel-3 WFR images. The two candidates are then compared to in-situ chl-a concentration samples, and for one which has a better accuracy in terms of root-mean-squared-error ($RMSE$) and coefficient of determination (R^2) is utilized in the first stage. As in the next stage, the in-situ chl-a concentration samples are classified into 2 groups, one for training samples and the other for testing samples. The in-situ training samples are used to train and refine the pre-trained model, and then the refined model is performed to the in-situ samples to check the performance. In addition, as the chl-a concentration from in-situ, IRTM-NN, and OC4Me have different range, a standardization by means of mean and standard deviation was performed to each.

Hyper-parameter adjustment. For training the proposed model, *Adam* optimization algorithm which derived from adaptive moment estimation is utilized for weight update. The algorithm combines the advantage of *AdaGrad* and *RMSProp* optimization algorithms by adapting the learning rates which based on both the average first moment and second moment of gradients (Kingma and Ba, 2015). It is able to handle large datasets and high-dimensional unknown parameter spaces with only little memory requirement. The parameters in the *Adam* optimization algorithm were tuned following the default in *Keras*. Moreover, mean-squared-error (MSE) was selected as the loss function. In addition to prevent the problem of overfitting, dropout and $L2$ regularization in kernel were used. The dropout removes out some neurons temporally when computing the loss function while $L2$ regularization penalizes larger weights by adding Frobenius norm to the loss function. The two regularizations were tuned respectively to 0.5 and 0.001 (learning rate).

4. RESULTS AND DISCUSSIONS

4.1. Existing Retrieval Model for Pre-training

The estimated chl-a concentration from IRTM-NN and OC4Me channels which collocated with in-situ stations were extracted and compared to in-situ chl-a concentration. The two channels performed slightly similar in each campaign, around 7 mg/L and 0.2 in terms of $RMSE$ and R^2 respectively, as shown in Table 3. Here, considering the $RMSE$, the estimation from IRTM-NN channel was selected and will be utilized in the pre-training part of the proposed two-stage training.

Table 2. Performance comparison of IRTM-NN and OC4Me channel

Campaign #	<i>RMSE</i>		<i>R</i> ²	
	IRTM-NN	OC4Me	IRTM-NN	OC4Me
1 st	5.5757	5.7412	0.0031	0.0827
2 nd	8.0890	8.2377	0.4320	0.3876
3 rd	10.0080	10.3608	0.2106	0.1405
4 th	6.5910	6.9167	0.2411	0.3480
5 th	7.2506	7.6134	0.3191	0.4984
6 th	6.1087	5.8565	0.0166	0.0878
7 th	8.4196	8.5385	0.3790	0.4966
Average	7.4347	7.6092	0.2288	0.2917

4.2. Proposed Model Performance

In this study, the datasets from dry season are used to training the proposed CNN model, which are from the 3rd to the 7th campaigns. Meanwhile, datasets from the other campaigns are used for testing. In this section, the proposed model is trained with the proposed two-stage training and compared to a training with one stage only, that is using estimation from IRTM-NN channel.

The performance evaluation reveals that the use of IRTM-NN channel in pre-training and in-situ chl-a concentration in refinement are able to result a more accurate chl-a concentration estimation. As shown in Table 3, the proposed model with two-stage training was applied, had a lower *RMSE* and higher *R*² value than with one-stage only was applied, both in training and testing datasets. This means that the two-stage training was able to enhance the estimation accuracy of chl-a concentration estimation.

Table 3. Performance comparison of proposed model with and without proposed two-stage

Campaign #	<i>RMSE</i>		<i>R</i> ²	
	Two-stage	One-stage	Two-stage	One-stage
1 st	2.6561	3.7352	0.0523	0.0305
2 nd	0.8772	1.3982	0.3907	0.3915
3 rd	1.9082	1.0453	0.0842	0.1308
4 th	0.2988	2.5705	0.6541	0.1167
5 th	0.8989	2.3315	0.5946	0.2538
6 th	2.1647	2.8652	0.1757	0.2719
7 th	0.6688	1.1856	0.5128	0.1125
Average	1.3533	2.1616	0.3521	0.1868

4.3. Proposed Model vs Existing Chl-a Retrievals

The proposed CNN model with two-stage training was applied, was compared to existing chl-a concentration estimation which corresponds to 3-band model from (Gitelson et al., 2008), 2-band model from (Moses et al., 2009), and normalized differentiate chlorophyll-a index from (Mishra and Mishra, 2012), where the regression coefficients of the three existing models were tuned according to the training datasets (see Table 4).

Table 4. Existing model and regression coefficients tuning

Model name	Chl-a model	a_1	a_0
3-band model	$[R_{rs}^{-1}(665) - R_{rs}^{-1}(709)] \times R_{rs}(754)$	2.5145	8.7236
2-band model	$[R_{rs}(709) \times R_{rs}^{-1}(665)]$	0.9880	7.7688
NDCI	$[R_{rs}(709) - R_{rs}(665)]/[R_{rs}(709) + R_{rs}(665)]$	-3.0256	8.7044

Table 5 summarizes the performance of the three existing chl-a concentration retrieval models. Compared to them, the proposed model is, in overall, able to estimate a better accuracy of chl-a concentration estimation, in terms of $RMSE$ and R^2 . However, the performance evaluation in the 1st and 6th campaign of the proposed model had lower $RMSE$ than the three existing retrievals. As for the 6th campaign evaluation, even the $RMSE$ of proposed model was lowest, the difference was insignificant. This means that the performance evaluation in this campaign was similar between the four. Hence in the 1st campaign evaluation, the in-situ chl-a concentration between this campaign dataset with training datasets was found to be different (see table 1). Broadly known that NN attempts to learn and to fit the training dataset. Thus, the different in term of range of in-situ chl-a concentration between the 1st campaign dataset and the training datasets might make the proposed model to retrieve the chl-a concentration unsuccessfully.

Table 5. Performance comparison of existing chl-a retrieval

Campaign #	$RMSE$			R^2		
	3-band	2-band	NDCI	3-band	2-band	NDCI
1 st	2.2423	2.2703	2.2866	0.3045	0.3173	0.3220
2 nd	0.9263	0.9191	0.8985	0.6013	0.5912	0.6121
3 rd	2.5942	2.5678	2.5900	0.0209	0.0160	0.0169
4 th	0.8801	0.9106	0.8810	0.0182	0.0059	0.0062
5 th	1.2107	1.2292	1.2093	0.6468	0.5761	0.5726
6 th	2.1129	2.1312	2.1619	0.0006	0.0001	0.0103
7 th	0.7858	0.8095	0.7843	0.3349	0.3980	0.4506
Average	1.5361	1.5482	1.5445	0.2753	0.2721	0.2844

5. CONCLUSION AND FUTURE WORKS

A CNN model for chl-a concentration retrieval was proposed in this study. The model contains three stage, that are band expansion, feature extraction, and chl-a concentration estimation. The three stage allowed the model to cover specific problems of statistical complexity, satellite error, and estimation, respectively. In addition, a two-stage training, taking the use of Sentinel-3 WFR channel, that is IRTM-NN channel, and in-situ chl-a concentration, was also proposed. The result revealed the capability of the proposed model with two-stage training in estimating the chl-a concentration in the seven campaign with a good accuracy in terms of $RMSE$ and R^2 . Moreover, the proposed model was compared to existing chl-a concentration estimation from (Gitelson et al., 2008), (Moses et al., 2009), and (Mishra and Mishra, 2012). This might be happened because of that the range of in-situ chl-a concentration between the 1st campaign dataset and the training datasets is different to each other, which may lead to an inconsistent estimation. Other than that, the proposed CNN model was able to outperform the three existing retrievals. In the future, a multiple output would be inserted into the model, where the output is not only chl-a concentration, but also other water quality parameters such as colored dissolved organic matter (CDOM), total suspended solid/matter (TSS/TSM), etc.

References

Barzegar, R., Skordas, K., Tziritis, E., Deo, R., Fijani, E., 2018. Design and implementation of a hybrid model based on two-layer decomposition method coupled with extreme learning

- machines to support real-time environmental monitoring of water quality parameters. *Sci. Total Environ.* 648, 839–853. <https://doi.org/10.1016/j.scitotenv.2018.08.221>
- Boyd, C.E., Tucker, C.S., 1998. *Pond Aquaculture Water Quality Management*.
- Bricaud, A., Morel, A., Babin, M., Allali, K., Claustre, H., 1998. Variations of light absorption by suspended particles with chlorophyll a concentration in oceanic (case 1) waters : Analysis and implications for bio-optical models. *J. Geophys. Res.* 103, 31,033-31,044.
- Buckton, D., O'Mongain, E., Danaher, S., 1999. The use of Neural Networks for the estimation of oceanic constituents based on the MERIS instrument. *Int. J. Remote Sens.* 20, 1841–1851. <https://doi.org/10.1080/014311699212515>
- Gitelson, A.A., Dall'Olmo, G., Moses, W., Rundquist, D.C., Barrow, T., Fisher, T.R., Gurlin, D., Holz, J., 2008. A simple semi-analytical model for remote estimation of chlorophyll-a in turbid waters: Validation. *Remote Sens. Environ.* 112, 3582–3593. <https://doi.org/10.1016/j.rse.2008.04.015>
- González Vilas, L., Spyrakos, E., Torres Palenzuela, J.M., 2011. Neural network estimation of chlorophyll a from MERIS full resolution data for the coastal waters of Galician rias (NW Spain). *Remote Sens. Environ.* 115, 524–535. <https://doi.org/10.1016/j.rse.2010.09.021>
- Hajigholizadeh, M., 2016. *Water Quality Modelling Using Multivariate Statistical Analysis and Remote Sensing in South. Florida International University*. <https://doi.org/10.25148/etd.FIDC001230>
- Hoagland, A.P., Anderson, D.M., Kaoru, Y., White, A.W., Hoagland, P., 2012. The Economic Effects of Harmful Algal Blooms in the United States : Estimates , Assessment Issues , and Information. *Assessment* 25, 819–837.
- Ioannou, I., Gilerson, A., Gross, B., Moshary, F., Ahmed, S., 2013. Deriving ocean color products using neural networks. *Remote Sens. Environ.* 134, 78–91. <https://doi.org/10.1016/j.rse.2013.02.015>
- Ioannou, I., Gilerson, A., Gross, B., Moshary, F., Ahmed, S., 2011. Neural network approach to retrieve the inherent optical properties of the ocean from observations of MODIS. *Appl. Opt.* 50, 3168. <https://doi.org/10.1364/ao.50.003168>
- Kingma, D.P., Ba, J., 2015. Adam: A Method for Stochastic Optimization, in: *ICLR*. pp. 1–15.
- Laguna Lake Development Authority, 2013. *Laguna de Bay 2013 Ecosystem Health Report Card 1–8*.
- Maruyama, S., 2017. Identification of waters incorporated in laguna lake, republic of the philippines, based on oxygen and hydrogen isotopic ratios. *Water (Switzerland)* 9. <https://doi.org/10.3390/w9050328>
- Mishra, S., Mishra, D.R., 2012. Normalized difference chlorophyll index: A novel model for remote estimation of chlorophyll-a concentration in turbid productive waters. *Remote Sens. Environ.* 117, 394–406. <https://doi.org/10.1016/j.rse.2011.10.016>
- Morel, A., Maritorena, S., 2001. Bio-optical properties of oceanic waters: A reappraisal. *J. Geophys. Res. Ocean.* 106, 7163–7180. <https://doi.org/10.1029/2000JC000319>
- Moses, W.J., Gitelson, A.A., Berdnikov, S., Povazhnyy, V., 2009. Satellite estimation of chlorophyll-a concentration using the red and NIR bands of MERIS - The azov sea case study. *IEEE Geosci. Remote Sens. Lett.* 6, 845–849. <https://doi.org/10.1109/LGRS.2009.2026657>
- Santos-Borja, A., Nepomuceno, D.N., 2006. Laguna de Bay: Institutional development and change for lake basin management. *Lakes Reserv. Res. Manag.* 11, 257–269. <https://doi.org/10.1111/j.1440-1770.2006.00310.x>
- Shumway, S.E., 1990. A Review of the Effects of Algal Blooms on Shellfish and Aquaculture. *J. World Aquac. Soc.* 21, 65–104.