

Remote sensing based indices for mapping of water quality in Thai Binh river at Hai Duong province

Nguyen Quoc Phi¹, Vu Manh Tuong², Nguyen Quang Minh³

¹Hanoi University of Mining and Geology (HUMG)

nguyenquocphi@humg.edu.vn

²Hai Duong Provincial Environmental Protection Department

thieudcxh@gmail.com

³Hanoi University of Mining and Geology (HUMG), Vietnam Academy of Sciences and Technology

hanoienvi@gmail.com

Abstract

This research explored the potential of remote sensing to estimate the relationship between satellite spectrals and riverine environments, especially during extreme events when routine in-situ measurements are not available. Water quality parameters of Turbidity, Total Suspended Solids (TSS), Total Dissolved Solids (TDS), Biochemical Oxygen Demand (BOD), Chemical Oxygen Demand (COD), Amoni concentration, Chlorophyll-a and the Trophic state index (TSI) were calculated using the Landsat 8 data and were correlated to the near in-situ measurements using regression equations for quantitative estimation of water quality in the session of Thai Binh River at Hai Duong province. This technique, using the obtained coefficients was applied to map the water quality in the Thai Binh River during the seasons of Sep. 2016, Nov. 2019 and Jun. 2020, and found in a good general agreement with the field measurements. The preliminary results indicate that (1) Spectral bands from Landsat images have the potential to map the spatial distribution of water quality in river environments, (2) Remote sensing data can be used for quantitative estimation of water quality parameters in river environments when coupled with linear regression equations, and (3) the same approach can be used to estimate water quality in rivers within reasonable error limits. Acquisition of more in-situ measurements of are on going to derive more general regression coefficients and achieve more validation results.

Keywords: RS-based indices, water quality mapping, Thai Binh river, Hai Duong province

1. Introduction

Water quality is one of the most important factors for riverine ecosystems. Water quality parameters of Turbidity, Total Suspended Solids (TSS), Total Dissolved Solids (TDS), Biochemical Oxygen Demand (BOD), Chemical Oxygen Demand (COD), Amoni concentration, Chlorophyll-a and the Trophic state index (TSI) in the water affect the river habitats strongly. These parameters reflect the physical and chemical property of water. They influence the total primary productivity by affecting the transmission of light in water, as well as the transition of heavy metal and the micro-

pollutant. Suspended sediment concentration and biochemical parameters are spatially inhomogeneous and their spatial distribution are difficult to measure with the routine in-situ monitoring method. Remote sensing is an efficient method, which can provide realistic water quality data with large spatial distributions for water resource study. In this study, remote sensing techniques are used for mapping of water quality in the Thai Binh River at Hai Duong province. Previous developed indices such as Turbidity Index, Chlorophyll-a Index (Frohn and Autrey, 2007; Bee, 2008), Normalized Difference Suspended Sediment Index (NDSSI) (Hossain và nnk, 2006) and Normalized Suspended Material Index - NSMI (Montalvo, 2010) were also used to test the applicability in the study area. The study finds that the use of previous approaches can be used with consideration of measurement data. The regression equations from remote sensing data are effective approaches to estimate the water quality in study area and they all have reasonable agreements with in-situ measurements.

Estimation of water quality from optical satellite data using regression equation was studied in many regions over the world. The available techniques can be categorized in four general groups: (1) simple regression (correlation between single band and in-situ measurements) (Williams and Grabau, 1973), (2) spectral unmixing techniques (Gomez et al., 1997), (3) Band ratio technique using two and more bands (Lathrop, 1992; Populus et al., 1995; Wang et al., 2000), and (4) multiple regressions (using multiple bands and in-situ measurements) (Binding et al., 2005).

Usually when suspended sediment concentrations are high, the backscatter/reflectivity of water is high. There are three matters dominate the reflectance of inland water, which are yellow substance, suspended sediment and phytoplankton. Yellow substance is a soluble matter, which has no scatter capability, but it has a strong absorption effect on short-wave bands that highly reduce the underwater downwelling irradiance. Therefore, when the absorption of water itself and yellow substance is small, actual suspended sediment information could be obtained (Wang et al., 2003).

In this research, various band math equations have been developed to determine many components of water quality, aiming to discover new mathematical relationship to identify and discriminate the suspended materials and biochemical concentrations in river water. The linear regression shows that the vast majority of image-estimated parameters are closely correlated with the in-situ measurements. It has been observed that although remote sensing has been considered as a proven technique for water quality estimation, all the developed models and algorithms are applicable for specific areas and environments. To address this issue in this research we attempted to explore the potential of remote sensing to develop regression equations that can be used in regional riverine environments like study area, especially during extreme events when routine in-situ measurements are not available.

2. Materials and methods

2.1. Data used

The data used for this study is mosaic images of the remotely sensed data of Landsat 8 (OLI). The system is designed to collect 15m resolution panchromatic data and different bands of data in the

visible, NIR, and MIR (Middle Infra Red) spectral regions at a resolution of 30m and 60/100m resolution data at Thermal Infra Red (TIR) band. Total 16 Landsat 8 OLI imagery from 2016 to 2020 were collected and the images acquired in 30/9/2016, 10/11/2019 and 21/6/2020 were used for further analysis. Figure below shows the nature of Landsat imagery prepared acquired over above 3 time periods.

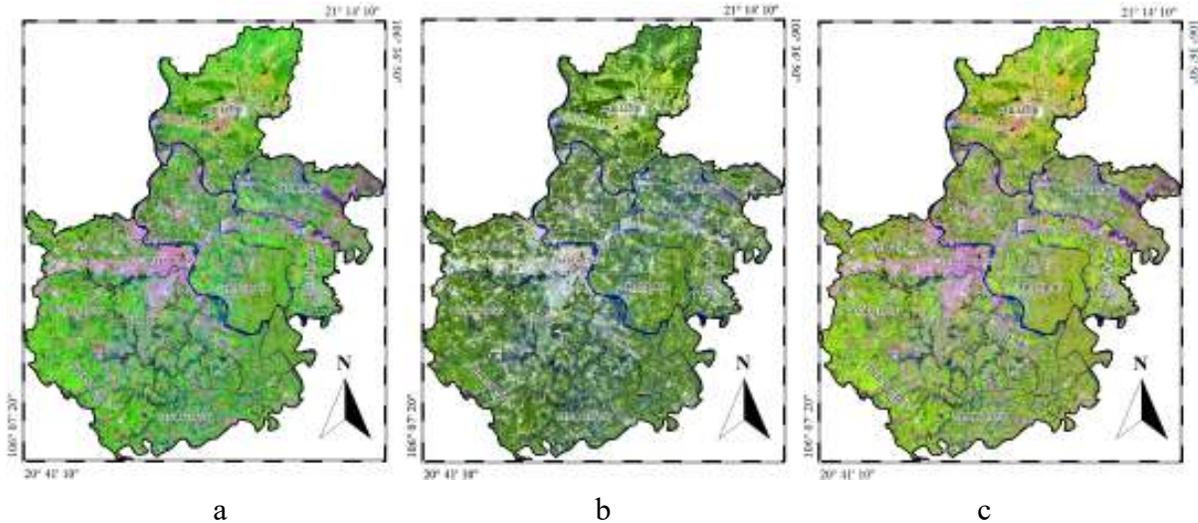


Figure 1. Landsat 8 images of study area on 9/2016 (a), 11/2019 (b) and 6/2020 (c)

To remove the systematic errors and improve the quality, DN values of Landsat image were converted to radiance. The conversion of DN values to radiance is based on a calibration curve of DN (Chander and Markham 2003; Negi et al. 2009). All the images were projected to the Universal Traverse Mercator (UTM) coordinates zone 48. The spheroid and datum was also referenced to WSG84. Enhancement of the images using histogram equalization techniques was later performed on all the images.

It has been proven that the satellite data and the ground measurement data should be taken almost simultaneously. Very fortunately, 23 of the measurements are found very close to the dates of the available Landsat 8 images of study area. Normally only the images within a 10-day overlap with the in situ data can be selected. There are a total of 23 ground sampling stations distributed on the river system.

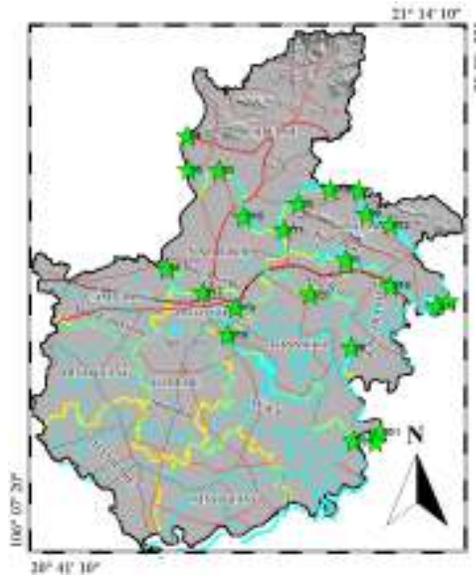


Figure 2. Location of 23 collect samples

2.2. Satellite-based water quality parameters

Landsat 8 OLI imagery were used to determine spatial distribution of the relative variation of water quality in the river environment. The water quality parameters of Turbidity, Total Suspended Solids (TSS), Total Dissolved Solids (TDS), Biochemical Oxygen Demand (BOD), Chemical Oxygen Demand (COD), Amoni concentration, Chlorophyll-a and the Trophic state index (TSI) values were correlated to in-situ measurements to estimate their concentration quantitatively.

Table 1. Part of the in-situ measurements database

KẾT QUẢ PHÂN TÍCH NƯỚC SÔNG TỰ NHIÊN- HỆ THỐNG SÔNG THÁI BÌNH

KH	Vị trí quan trắc	Tọa độ	Tọa độ	Lưu lượng m ³ /s	Nhiệt độ	pH	Độ dẫn µS/cm	Độ đục NTU	TDS mg/l	Mnói 0/∞	DO mg/l	F ⁺ mg/l
		X	Y									
S1	Sông Thương, cách ngã ba sông Lục Nam và sông Thương 500m về phía thượng lưu	21°16'13"	106°31'85"	385,20	26,8	7,00	196	60	98	0,10	6,49	<0,30
S2	Sông Lục Nam, cách ngã ba sông Lục Nam và sông Thương 500m về phía thượng lưu	21°16'14"	106°32'10"	424,32	26,4	7,20	202	66	101	0,10	5,72	<0,30
S3	Sông Thương, cách đền Kiếp Bạc	21°14'66"	106°32'46"	375,90	27,7	7,30	196	57	98	0,10	5,85	<0,30
S4	Sông Cầu, cách điểm nối giữa sông Cầu và sông Thương 500m về phía thượng lưu	21°07'31"	106°17'46"	277,20	31,8	6,92	174	65	87	0,08	4,00	0,32
S5	Sông Đường, cách điểm nối giữa sông Đường và sông Thương 500m về phía thượng lưu	21°06'28"	106°16'56"	789,25	29,8	7,28	178	64	89	0,09	6,23	<0,30
S6	Sông Thái Bình tại xã Nhân Thọ, cách ngã ba sông Thái Bình và sông Kinh Thầy 500m về phía thượng lưu	21°30'10"	106°21'36"	396,50	29,2	7,20	190	47	95	0,10	5,74	<0,30
S7	Sông Thái Bình tại xã Thái Tân huyện Nam Sách	20°59'4"	106°16'39"	182,40	31,7	7,20	162	49	81	0,08	5,93	0,34
S8	Sông Thái Bình tại xã An Bình, huyện Nam Sách (sông Kinh Thầy)	21°35'00"	106°22'3"	936,00	28,2	7,15	160	134	80	0,08	5,48	<0,30
S9	Sông Đông Mai tại phường Văn Đức, TP. Chí Linh	21°07'85"	106°44'07"	34,50	30,6	7,02	162	126	81	0,08	4,85	<0,30
S10	Sông Đà Vách, xóm Trại khu Từ Lạc (gần khu vực chế biến than)	21°22'20"	106°35'47"	272,40	29,3	7,6	254	65	127	0,13	5,83	<0,30
S11	Sông Đà Vách, gần trạm cấp nước Tân Dân, Kinh Môn	21°33'10"	106°34'41"	402,48	30,8	7	212	72	106	0,11	5,79	<0,30
QCVN 08-MT:2015/BTNMT (Mức B1)		-	-	-	-	5,5-9	-	-	-	-	24	1,5

It has been observed for Landsat OLI imagery that Blue band and Near-infrared band are most sensitive to water and water transparency (turbidity). These bands usually gives the highest and lowest reflectance values respectively for water. These characteristics have been observed for water with different levels of turbidity. Biochemical parameters were also sensitive to the Blue, Green, Red and Near infrared bands. According to Nguyen et al., (2016) for any vegetation band 5 (near-infrared) and band 4 (red) of Landsat 8 OLI data always gives the highest and lowest reflectance respectively.

Near real time of the corresponding image acquisition dates in-situ measurements of water quality in the Thai Binh River were obtained from 23 stations. To determine the most suitable coefficients

to estimate water quality in the river water using remote sensing data, the obtained values were plotted against the corresponding near real time in-situ measurements. The relationship between calculated values and in-situ measurements were interpreted using different numerical equations including linear, exponential, logarithmic, polynomial and power function. The coefficients associated with the equation that achieved highest correlation coefficient (R^2) value were considered suitable to use water quality estimation in the study area. Accordingly the polynomial equations was found to have R^2 values of over 0.8 and considered the most suitable equations initially. After a careful observation it was noticed that the polynomial equations have a potential to provide accurate estimation of TSS and TDS concentrations in low turbid and high turbid water because they have the best fit. However, the curves are very flat in the low turbid region and seem not to be capable of detect variation in the estimation. More importantly, the trend of increase in concentration at high turbidity values may result in unphysical predictions. The linear equations, although posses lower R^2 values than the polynomial equations, they show the potential to detect the variability in suspended materials and biochemical concentrations in both low and high turbid water. Therefore, the power equations were considered to be the most suitable approach to estimate the suspended materials and biochemical concentrations in the study area.

3. Results and discussion

3.1. Regression analysis

Landsat 8 spectral bands were converted to DN and then, these values are combined and calculated based on Equation (1), which is the general predictive equation for water quality parameter, and then are compared with the ground measurements.

$$y = ax + b \quad (1)$$

Where, x is the DN values and y stands for the individual water quality parameter.

a and b are the coefficients fitting to the in-situ data by the regression analysis.

Different models are used for the estimations of 8 water quality parameters by utilizing linear regression analysis. Comparisons are conducted for individual band or band ratio, while the combinations of reflectance value are more complex which include $TM3*TM4$ (Chen et al., 1996), $TM3*TM4/\ln TM1$ (Li et al., 2007), $TM3*TM4/\ln(TM1+TM2)$ (Li et al., 2007), $TM3/TM4$, $(TM4-TM3)/(TM4+TM3)$, $TM4/TM1$, $TM1/TM2$ (Dwivedi and Narain, 1987), $\ln(TM1/TM2)$ (Li et al., 2007), and $(TM1-TM3)/(TM2-TM3)$. Once the correlation coefficient R^2 is considered as high enough, a regression equation will be calculated and this equation can be used for the whole image and generates a water quality maps.

By applying the regression equations to each of the images in the database, the spatial patterns, as well as the temporal trends of 8 water quality parameters, can be investigated.

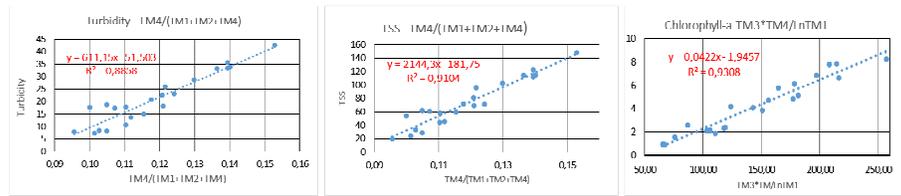


Figure 3. Scatter plots comparisons for water quality estimations

The comparison results for the estimated values and in-situ measurements of different parameters. The prediction made by the linear models have a very good relationship with the suspended materials and biochemical data, and the scatter points on the plots seem to be close to the line, which means their relationships are probably linear. All the datasets show good agreements, with the R^2 over 0.5. Therefore, it can be concluded that the Landsat 8 OLI images are reliable enough to predict the water quality in Thai Binh river system at Hai Duong province.

3.2. Spatial and temporal analysis

The seasonal patterns and temporal trends of water quality in Lake Simcoe are both analyzed in this study, as the available data for Thai Binh river system have a long time span of 4 years from 2016 to 2020. The spatial patterns of all 8 parameters are displayed in terms of the concentration maps.

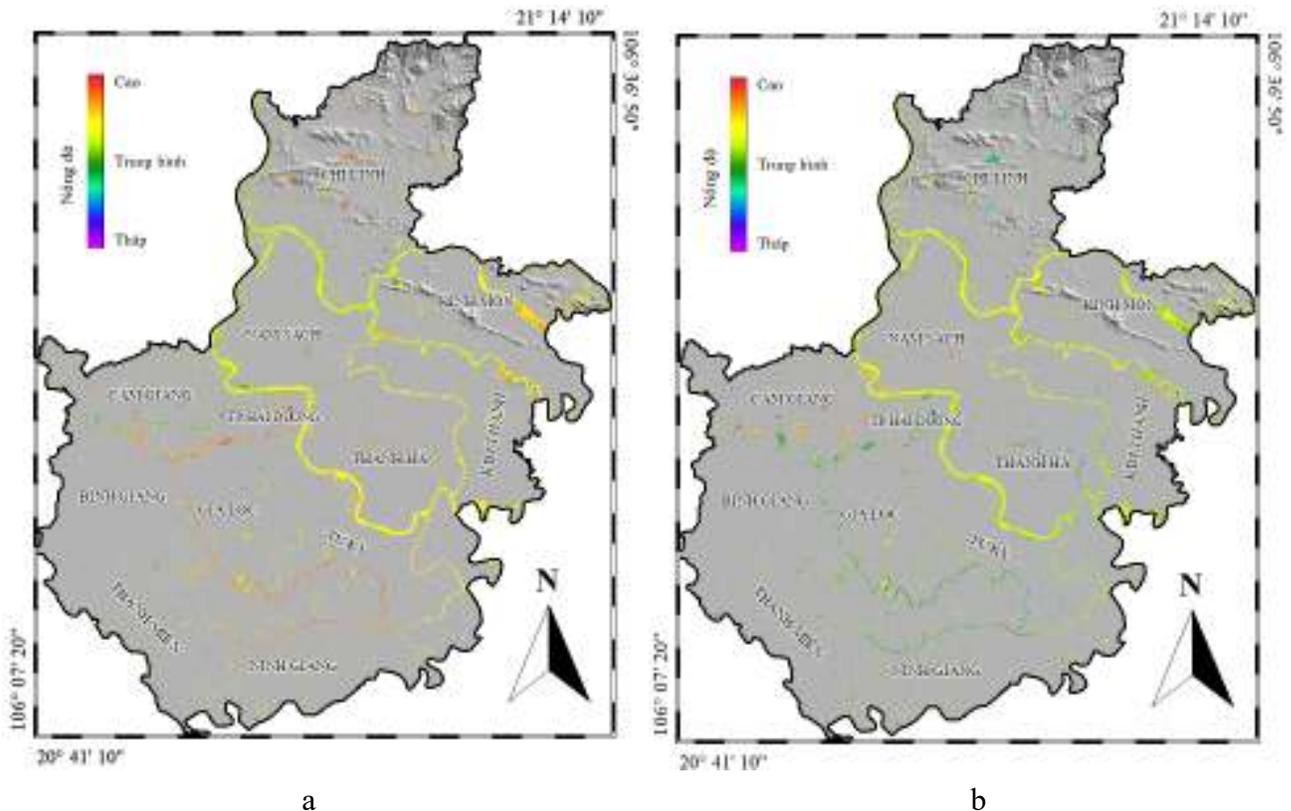


Figure 4. Turbidity (a) and TSS (b) maps in September 2016

From the above maps, it is obvious that the upper and lower sections of the river is normally clearer than the middle area, and the turbid area is concentrated near the Hai Duong city as well as industrial zones (An Phat, Viet Hoa - Kenmark, Nam Sach, Dai An, Lai Vu, Lai Cach, etc.)

The BOD₅, COD, NO₃⁻ are clearly show both the effects population density at major cities and towns in their northern part and agricultural activities in the southern part of study area. The amount of fertilizer is derived from the land use, which is predominately agricultural land, and in addition, the vegetation lands occupy about 67% of the terrestrial area. Hai Duong is one of the provinces with a very high urbanization rate in recent years. About 15% of the land area is urban, where the population has doubled in the past twenty years.

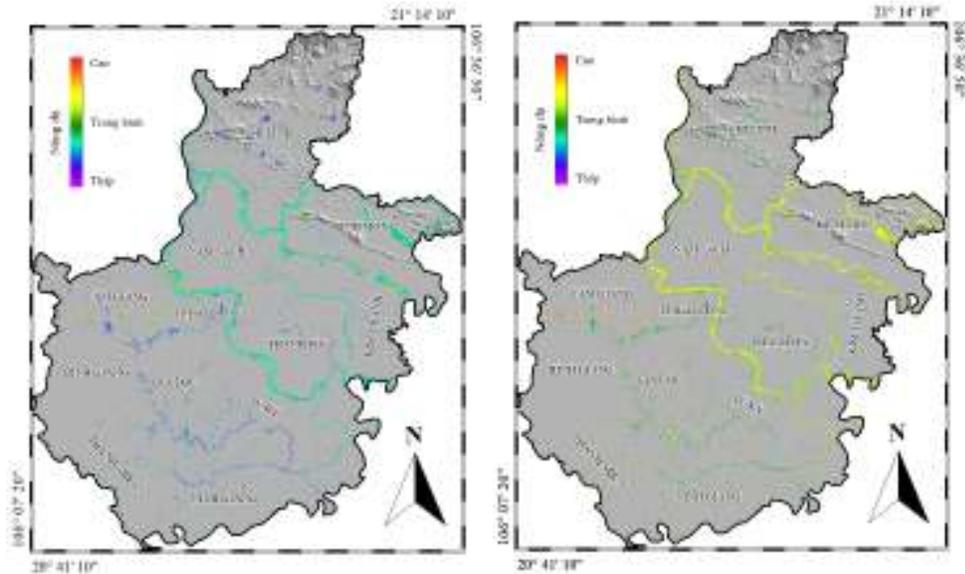


Figure 5. Calculation results of BOD₅ content in September 2016 (a) and November 2019 (b)

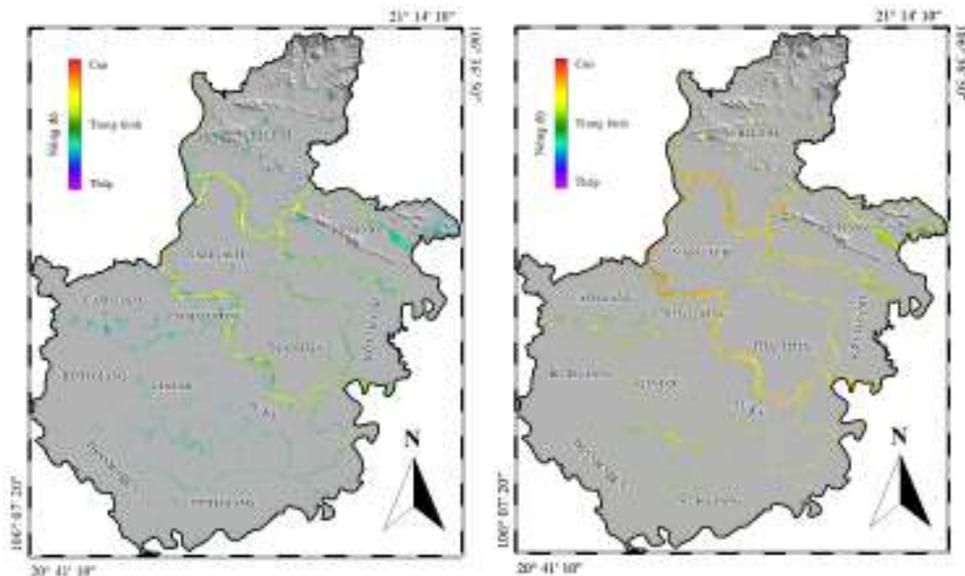


Figure 6. Calculation results of COD content in September 2016 (a) and November 2019 (b)

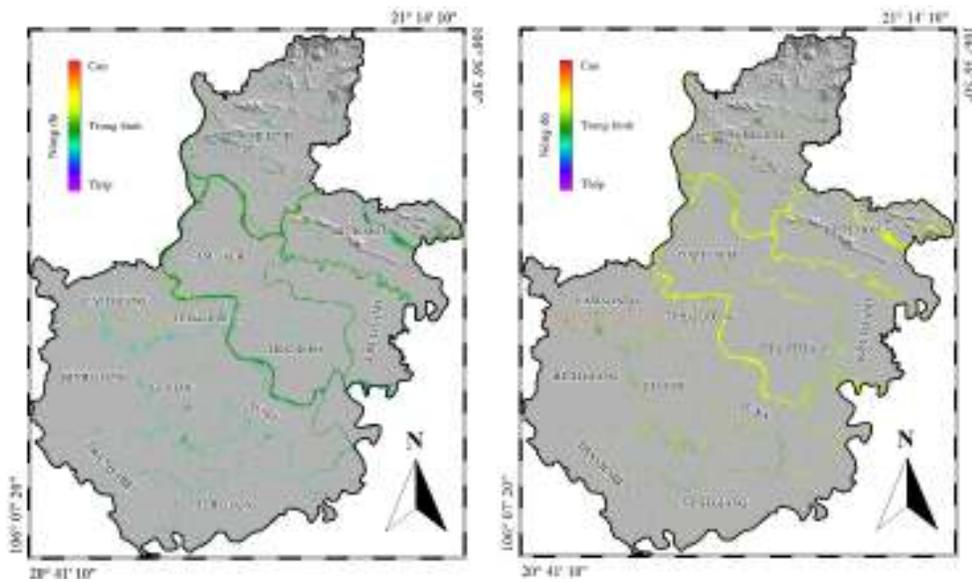


Figure 7. Calculation results of NO_3^- content in September 2016 (a) and November 2019 (b)

Correlation coefficients of BOD_5 , COD and NO_3^- are somehow lower than the parameters of suspended materials, with the values of 0.348, 0.439 and 0.679, respectively. The distribution of population, industrial activities, the landforms and land uses in the area significantly influence the water flowing to the Thai Binh river system, and thus the water quality of the water. All of these factors can contribute to the instability of river water and the low agreements between the satellite-estimated parameters and the in-situ measurements.

For the parameters of suspended materials (turbidity, TSS and TDS), it clearly shows that in the late-summer period, the river usually has its worst water clarity, there are more lights and signals reflected to the remote sensor. Therefore, the relationships between the reflectance values and in-situ data become stronger.

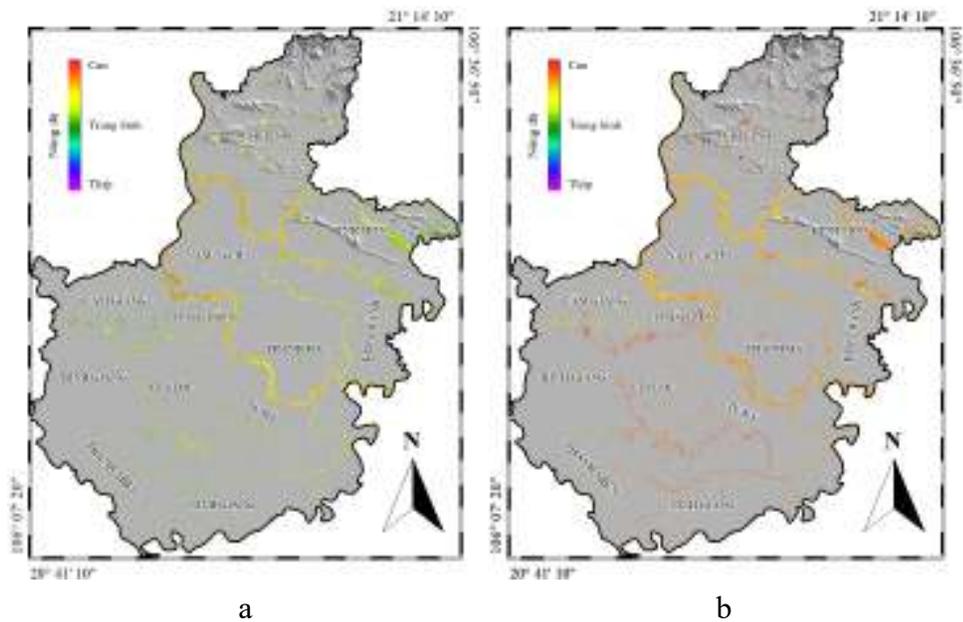


Figure 8. Calculation results of TDS content in September 2016 (a) and November 2019 (b)

There might be some other components in the water such as the garbage left by the residents. The eroded soil will make the water turbid as well.

In the scale of the whole province, the northwest part of the study area has worse water quality than the southeast part. This indicates that the northwest part is suffering from a more severe water quality problem than the southeast part. It can also be found that the water quality is turning bad gradually following the direction from southwest to northeast.

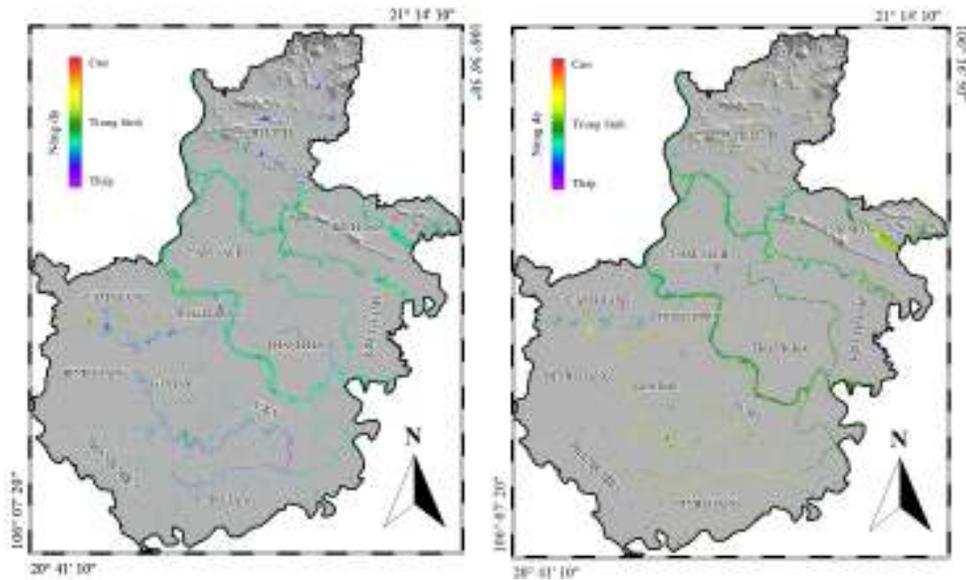


Figure 9. Calculation results of Chlorophyll-a (a) and Trophic State Index (b) in study area

The study area is also suffering from the most severe environmental problem with agricultural activities. The concentration map estimated by the regression equations can indicate the spatial distribution of the eutrophic areas. The contribution of TP, TN and chl-a are relatively high in drainage water pumped from the intensively cultivated land. However, the trophic state is still not too high in the river compare to the lakes in the area. The Carlson's modified trophic state index (TSI_m) is applied in this study, as it only uses the chl-a concentration to represent the trophic state index of both the river system and inland lakes in the area. The trophic state on Thai Binh river system show oligotrophic state, it generally host very little or no aquatic vegetation and are relatively clear, while eutrophic lakes tend to have large quantities of organisms, including algal blooms. In some cases, the algal biomass in lakes reaches too high and massive fish die-offs occurred as decomposing biomass deoxygenates the water.

The colour in the water can increase the absorption of light and therefore decrease the signal back to the remote sensors. The atmospheric conditions (e.g., haze and water vapor content) clearly affect the light reflected by land and water surfaces as it travels back toward the satellite sensors.

The trends of all suspended sediment concentration and biochemical parameters imply that the water quality of the entire area has worsen, especially the water quality in the northern area where the population density is high and the concentration of industrial zones, and more attention needs to be paid to this area.

4. Conclusion

Parameters of suspended materials and biochemical concentrations were calculated using the Landsat 8 OLI imagery to map the spatial distribution water quality in the Thai Binh River at different time periods. This study builds the empirical relationships between the Landsat 8 images and the in-situ water quality data. Obtained values from regression equations were correlated with the near real-time in-situ measurements and the relationship between them were expressed by linear equations were found most suitable for quantitative estimation of water quality in the Thai Binh River. Obtained regression coefficients were applied to map the 8 water quality parameters in study area. The results were compared by the near real time measurement data and found in good general agreement qualitatively and quantitatively.

This research is still evolving and these initial results indicate that remote sensing based data has the potential to estimate relative variation of water quality and map the spatial distribution of suspended materials and biochemical concentrations in river environments in study area. Acquisition of more in-situ measurements of water parameters are on going to derive more general regression coefficients and achieve more validation results. However, there are usually a confusion of the pixels in satellite imagery between the water and soil if the spatial resolution of the satellite images is not high enough. Sometimes even if the stations are located perfectly on the imagery, it still may turn out that the DN values of soil rather than water are extracted. All of these reasons may more or less affect the results of estimation.

5. Reference

- [1].Aizaki, M., T. Iwakuma and N. Takamura, 1981. Application of modified Carlson's trophic state index to Japanese lakes and its relationship to other parameters related to trophic state, Research Report on National Institute of Environmental Studies, 23: 13-31.
- [2].Albright, M.F., 1996. Hydrological and nutrient budget for otsego Lake, N.Y. and relationships between land form/use and export rates of its sub-basins. Occas. Pap. #29. SUNY Oneonta Bio. Fld. SUNY Oneonta.
- [3].Binding, C.E., J.H. Jerome, R.P. Bukata, and W.G. Booty, 2007. Trends in water clarity of the lower Great Lakes from remotely sensed aquatic color, Journal of Great Lakes Research, 33(4):828-841.
- [4].Brando, V.E., and A.G. Dekker, 2003. Satellite hyperspectral remote sensing for estimating estuarine and coastal water quality, IEEE Transactions on Geoscience and Remote Sensing, 41(6): 1378-1387
- [5].Carlson, R.E. and J. Simpson, 1996. A Coordinator's Guide to Volunteer Lake Monitoring Methods. North American Lake Management Society. 96 pp.
- [6].Chen, C.Q., P. Shi, A and Q.P. Mao, 1996. Study on modeling chlorophyll concentration of surface coastal water using TM data, J. Remote Sensing, 11(3): 168-175.
- [7].Chen, Q., and Y. Zhang, 2007. Water quality monitoring using remote sensing in support of the EU water framework directive (WFD): A case study in the Gulf of Finland, Environment

- Monitoring and Assessment, 124(1-3): 157-166.
- [8]. Cox, R. M., R. D. Forsythe, G. E. Vaughan, and L. L. Olmsted, 1998. Assessing water quality in the Catawba River reservoirs using Landsat Thematic Mapper satellite data, *Lake and Reservoir Management*, 14: 405– 416.
- [9]. Dekker, A.G., R.J. Vos and S.W.M. Peters, 2001. Comparison of remote sensing data, model results and in situ data for total suspended matter (TSM) in southern Frisian lakes, *The Science of the Total Environment*, 268(1-3): 197-214.
- [10]. Dekker, A.G. and S.W.M. Peters, 1993. The use of the Thematic Mapper for the analysis of eutrophic lakes: a case study in the Netherlands, *International Journal of Remote Sensing*, 14(5):799-821.
- [11]. Fraser, R.N., 1998. Multispectral remote sensing of turbidity among Nebraska Sand Hills Lakes, *International Journal of Remote Sensing*, 19(15): 3011-3016.
- [12]. Harrington, J.A., F.N. Joe, Jr. and F.R. Schiebe, 1992. Remote sensing of Lake Chicot, Arkansas: Monitoring suspended sediments, turbidity, and Secchi depth with Landsat MSS data, *Remote Sensing of Environment*, 39(1): 15-27.
- [13]. Heiskary, S., J. Lindbloom and C.B. Wilson, 1994. Detecting water quality trends with citizen volunteer data, *Journal of Lake and Reservoir Management*, 9(1): 4-9.
- [14]. Hirthle., H. and A. Rencz, 2003. The relation between spectral reflectance and dissolved organic carbon in lake water: Kejimikujik National Park, Nova Scotia, Canada, *International Journal of Remote Sensing*, 24(5): 953-967.
- [15]. Trainer, T.J. and F.K. Sun, 1991. Image resampling in remote sensing and image visualization applications, *Proceeding of the SPIE's 1991 International Symposium on Optical Applied Science and Engineering*, San Diego, CA (USA), 21-26 July 1991.
- [16]. Ucuncuoglu, E., O. Arli, and A.H. Eronat, 2006. Evaluating the impact of coastal land uses on water-clarity conditions from Landsat TM/ETM+ imagery: Candarli Bay, Aegean Sea, *International Journal of Remote Sensing*, 27(17):3627-3643.
- [17]. Winter, J.G., M.C. Eimers, P.J. Dillon, L.D. Scott, W.A. Scheider, and C.C. Willox, 2007. Phosphorus inputs to Lake Simcoe from 1990 to 2003: Declines in tributary loads and observations on lake water quality, *Journal of Great Lakes Research*, 33(2): 381-396.



Letter of Acceptance

Dated: 23/05/2022

Dear Authors,

Based on the recommendations from the peer review board, I am delighted to inform you that your following manuscript has been accepted for publication.

Paper ID: HIW1-6T04
Title: Remote sensing based indices for mapping of water quality in Thai Binh river at Hai Duong province
Authors: Nguyen Quoc Phi
Vu Manh Tuong
Nguyen Quang Minh
Article received: 16/03/2022
Accepted on: 20/05/2022

Accepted papers will be published in the Proceeding Book of *"Innovative Water Solutions for Vietnam and Region"*.

Thank you very much for your contribution to the VACI 2022!

Best regards,

A handwritten signature in black ink, appearing to read 'Bui Du Duong'.

Dr. Bui Du Duong

Organizing Committee

VACI 2022 Vietnam International Water Week

Tel: (+84) 96-621-4316

Website: <http://vaci.org.vn>; <http://viww.vaci.org.vn>