

Study of NDVI Application on Turbidity in Reservoirs

Wen-Huan Chien¹ · Tai-Sheng Wang^{1,3} · Hui-Chung Yeh² · Tsu-Kuang Hsieh²

Received: 28 February 2015 / Accepted: 9 December 2015 / Published online: 24 February 2016
© Indian Society of Remote Sensing 2016

Abstract Turbidity (TB), an important factor that determines water quality, can influence the water supply of reservoirs. Using a submerged optical sensor, the TB of a body of water can be observed. This study aims to use remote sensing (RS) technology in order to examine TB variations of the water in reservoirs. In general, previous studies have focused on evaluating the relationships among a number of single-spectrum bands and in situ data. In this study, we included the Normalized Difference Vegetation Index (NDVI) among the TB monitoring methods for reservoirs and investigated its practical utility. The ratio of NDVI is between -1 and 1 . A negative value indicates water or snow; a value of 0 indicates rock or bare soil; and a positive value indicates a vegetative cover. We consider the value of NDVI to be close to -1 in clean water and close to 0 in turbid water. TB, which is the scattering degree of incident light into water, is observed using a submerged optical sensor. Water appears turbid if floating material or suspended solids are found. Therefore, we use NDVI in this study in an attempt to estimate TB concentration. Images from the Landsat-7 ETM+ satellite were used to observe several important reservoirs in northern Taiwan, and multiple linear regression (MLR) was used for analysis. This study examines the data of 47 samples and found the gaps between

the 3 days of the in situ data and the date the images were taken. We found that the NDVI has a negative correlation with TB. After including the NDVI in our model, its explaining ability and improvement rate increased by 11.2 and 8.72 %, respectively. Therefore, using the NDVI can provide additional reflective information, as well as improve the model's accuracy.

Keywords Turbidity · Remote sensing · Normalized Difference Vegetation Index · Multiple linear regression

Introduction

Turbidity (Turb / TB) is defined as the scattering degree of incident light into water. It is often used to determine water quality as water appears turbid when floating materials or suspended solids are present. A high TB level inhibits light transmission, thus affecting aquatic plants' photosynthesis and fish's ability to breathe; degermation for the purposes of water purification is also upset. In most cases, TB is observed using a submerged optical sensor (Garaba et al. 2014).

First developed in the 1960s, remote sensing technology has become a useful and comprehensive method of observation that has been applied to modern physics, space technology, electronic technology, computing technology, information science, and environmental science (Chang et al. 2012; Liou et al. 2012; Liou et al. 2015). Remote sensing technology provides the key for observing the earth and its environments (Liou et al. 2013). Because of its advantage of being able to evaluate environmental changes over large areas, remotely sensed images have been widely applied to obtain a variety of information about the earth's surface, ranging from data for military applications to environmental change in the

✉ Tai-Sheng Wang
m9204051@gmail.com

¹ Agricultural Engineering Research Center, Taiwan, Republic of China

² Department of Natural Resources, Chinese Culture University, Taiwan, Republic of China

³ Engineering Division, Agricultural Engineering Research Center, No. 196-1, Chungyuan Rd., Chungli District, Taoyuan City, Taiwan, Republic of China

vegetation cover, water pollution, and polar ice fluctuations (Liou et al. 2012).

In the light scattering phenomenon, light is degraded, and thus Remote Sensing (RS) technology can be applied to measure the concentration change of TB, an indicator of water quality. Typically, the unit used to report the measure of TB is the Nephelometric Turbidity Unit (NTU). RS technology is characterized by a large observation range and real-time functioning that returns quick and comprehensive monitoring results in order to increase the reliability of current the scientific monitoring of local water quality (Changliang 2013). Dogliotti et al. (2015) have indicated that ocean color remote sensing can be useful for mapping TB and the concentration of suspended particulate matter (SPM) in coastal waters. Koponen et al. (2001) and Wu et al. (2007) effectively used Landsat TM satellite images to perform statistical regression analysis using multiple single-bands to monitor the TB changes in Finnish Lake and Poyang Lake, respectively. Zhang et al. (2003) and He et al. (2008) also used Landsat TM satellite images to monitor the TB changes in Finland Gulf and Guanting Reservoir, respectively, and in addition to using multiple single-bands, these studies also used simple multi-band combination types as factors. Statistical regression was applied to create a prediction model in the above studies, wherein the coefficient of determination (R^2 or R squared), a number that reflects how well data fit a statistical model, was 0.709 and 0.878, respectively. These results demonstrate that model prediction accuracy increases with the inclusion of multi-band combination types. Furthermore, Zhang et al. (2003) used the Artificial Neural Network (ANN) for analysis and found improved accuracy of the prediction model, with the explanatory power of R^2 reaching 0.942. Dhillon and Mishra (2013) found that Landsat 7 ETM+ can be effectively used to estimate the Trophic State Index for Sukhna Lake. Garaba et al. (2014) relied on time series measurements at the Wadden Sea time series station Spiekeroog (WSS) in the southern part of the North Sea to empirically develop approaches for evaluating turbidity using ocean color remote sensing products (OCPs). The empirical approaches that they developed had good least squares linear correlations and statistical significance ($R^2 > 0.7$, $p < 0.001$). These OCP approaches enjoyed relatively low uncertainties when predicting turbidity with an encouraging mean absolute percent difference of less than 31 %.

Meanwhile, the single-band trial and error combination generates combined band type factors that significantly correlate to TB, thus increasing the accuracy of the prediction model; however, the feasibility of the combined type generated may not be common. Therefore, our aim is to integrate an index variable to interpret vegetation, which is the Normalized Difference Vegetation Index (NDVI), into the research to determine whether it can evaluate and improve the accuracy of TB concentration estimates.

Methodology

Normalized Difference Vegetation Index (NDVI)

In general, image classification is a part of the research that involves interpretation of the Normalized Difference Vegetation Index (NDVI). Equation 1 shows that the ratio of Near Infrared Red (NIR) to Red is between -1 and 1 . A negative value indicates water or snow; a value of 0 indicates rock or raw soil; and a positive value indicates a vegetative cover. Compared to other vegetation indexes, this index value range is considerably stable. Regarding multi-period Landsat-7 ETM+ satellite images, in addition to using NIR and Red, water messages that can be interpreted via NDVI are also included as factors to determine its potential for monitoring TB changes in reservoirs. The aim is to increase the prediction model's accuracy and establish common RS technology factors.

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad (1)$$

Where, NIR represents the short wave near the infrared red band, and Red represents the red band.

Image Preprocessing

The satellite images may have inaccuracies related to radiation measurements due to atmospheric scattering and sunlight incident angle and geometry due to distortion or deformation due to vehicle forms, orbital parameters of aviation, or surface curvature. Therefore, the data obtained from sensors may not correctly match clutter reflections. Furthermore, the ground's sample position may not precisely match the image's location coordinates on the map, which is known as the so-called "image distortion" problem (Liou et al. 2015). Therefore, the images obtained have been corrected prior to application to be classified. This process is referred to as "image correction" or "image processing" and includes geometric correction and radiometric correction (Wang 2010).

Multiple Linear Regressions (MLR)

Regression analysis (RA) is commonly used to comprehend the quantitative relationship between an independent variable and dependent variables. This statistical method is applied to determine the mutual correlation between two or more variables. That is, predictor x is used to explain or predict criterion y , the result of which is expressed in the relational function $y = f(x)$ by means of the least error square sum principle. New information with value x can be inserted into the relational function to predict the consequent y value. Multiple linear regression (MLR) analysis is often applied in order to express

dependent variable change. When working with regression analysis that has two or more variables, a regression surface is shown. MLR is represented with the following equation:

$$\hat{Y} = a_0 + a_1x_1 + a_2x_2 + \varepsilon \tag{2}$$

Where Y^{\wedge} is a dependent variable; a_0 is the intercept; a_1 and a_2 are independent variable coefficients; x_1 and x_2 are independent variables; and ε is the error. Therefore, Band-NIR, Band-Red of Landsat-7 ETM+ satellite images and the NDVI index are all used for analysis. Furthermore, the 47 data samples were divided into 38 training samples and 9 test samples, as shown in Table 1. MLR analysis is applied to develop a TB prediction model in order to study the influence on the model before and after including NDVI.

Case Study

Study Area

With regard to the studied areas, this study investigated northern Taiwan’s primary reservoirs, which are Feitsui Reservoir, Shihman Reservoir, Baoshan Reservoir, Yeonghershans Reservoir, and Mindah Reservoir, as shown in Fig. 1, among which Baoshan Reservoir and Yeonghershans Reservoir are considered off-channel reservoirs, which are water supply lakes built next to or close by a river, while the other reservoirs are considered on-channel reservoirs, which means that the dam site is actually located on the river.

In Situ Data

The in situ data of each of the reservoirs have been obtained through fixed samples located at WGS84 on the earth using

Global Positioning System (GPS). However, the in situ sampling periods for each reservoir differs. While the Reservoir Authority uses monthly periods, the Environmental Protection Administration (EPA) uses quarterly periods. In order to comply with satellite image dates, this research uses both monthly and quarterly sampling data to search the available satellite image data for different periods. Since the in situ historical data of observation cannot be easily matched to the satellite images of the same date, we chose to use satellite images (without clouds or rain) for the 3 days of gaps between the in situ data and the recording data. Therefore, this study consists of the search of 47 samples’ data

Image Data

This research features Landsat-7 ETM+ satellite images obtained from the Earth Resources Observation and Science (EROS) Center and taken by the Agricultural Engineering Research Center (AERC). An instrument carried onboard the Landsat-7 satellite malfunctioned, so all of the images received by the ETM+ Sensor were affected by the Scan Line Corrector being in the off (SLC-off) mode (USGS 2007), which generates noise, since 2003. Regarding imagery correction, only radiometric correction has been performed on the captured images. With regard to geometric correction, both coordinate conversion and image resampling have been performed in this research. To recover the Digital Number (DN) values of the image data for the corresponding water quality samples, the mean DN value of 3×3 Pixels on images is used to describe the corresponding ground points. Landsat-7 ETM+ has a total of seven spectrum bands that range from 0.45 to 2.35 μm to provide surface messages using visible light, NIR light, shore wave IR light, and thermal IR light, as shown in Table 2. The temporal resolution, also known as

Table 1 Satellite images and in situ samples (a total of 47 samples)

Types	Image Date	Sampling Date	Gaps of Days	Samples	Reservoirs
Training Set	1999/08/08	1999/08/09	1	5	Feitsui
		1999/08/10	2	4	Baoshan
				3	Yeonghershans
	2003/11/07	2003/11/10	3	6	Feitsui
				4	Baoshan
	2004/02/11	2004/02/09	2	7	Feitsui
				8	Shihman
				1	Baoshan
				38	Total
	Testing Set	2004/11/09	2004/11/08	1	7
2004/11/11			3	2	Yeonghershans
				9	Total

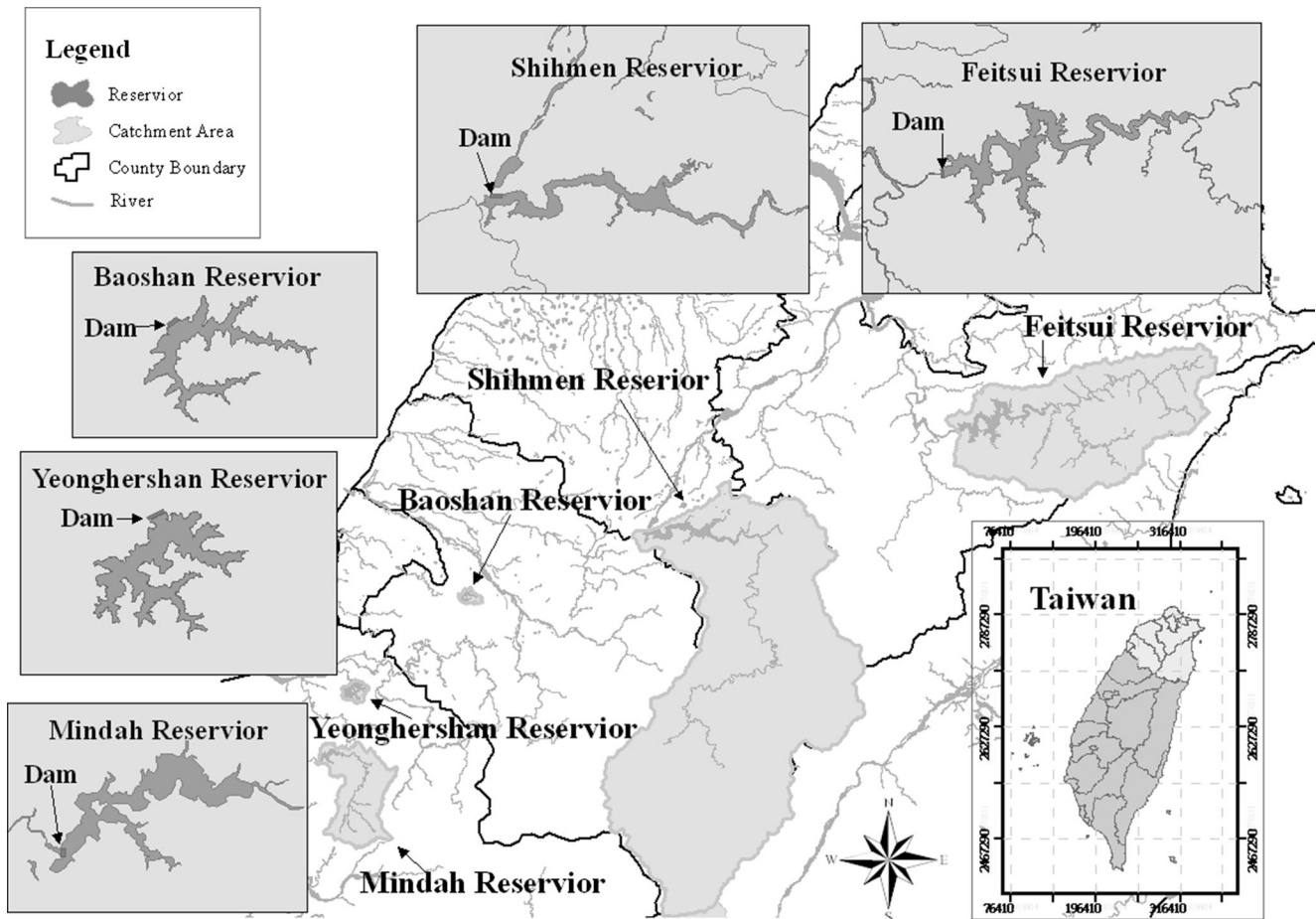


Fig. 1 Reservoir sites in northern Taiwan

the revisiting rate, is 16 days. Consequently, only two bands, Red and NIR, are used as investigative factors in this study in order to evaluate the possibility of using NDVI as a method for monitoring the TB change in reservoirs.

The spatial distribution of TB on images that cover the entire body of water in the reservoirs is drawn for comparison, as seen in the scatter plot in Fig. 2. However, as in the radiation image, the satellite images were distorted or deformed by geometric errors due to atmospheric scattering, sun angle difference, vehicle forms, surface curvature of the flight trajectory parameters, and other factors. Therefore, we performed both radiometric correction and geometric correction during image preprocessing to obtain accurate data and information.

Results

Data Pre-Processing

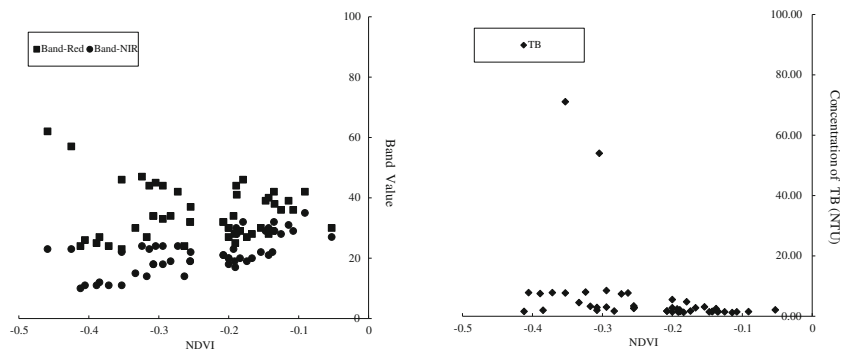
Regarding in-situ sampling time, this research included images within 3 days before and after the satellite

photography was taken. We used a total of six multi-period Landsat-7 ETM+ images, which is equivalent to 47 samples, as shown in Table 1. The two most recent images, which were taken in 2004, are used as test examples and consist of a total of nine samples, while the other images, which consist of a total of 38 samples, are used as training examples. The range of data was determined by matching, retrieving, and performing statistical analysis of

Table 2 Landsat-7 ETM+ image bands

Band	Spectral Resolution (μm)	Spatial Resolution (m)
B1 (Blue)	0.45 ~ 0.52	30
B2 Green)	0.52 ~ 0.61	
B3 (Red)	0.63 ~ 0.69	
B4 (NIR)	0.75 ~ 0.90	
B5 (SWIR)	1.55 ~ 1.75	
B6 (TIR)	10.4 ~ 12.5	60
B7 (SWIR)	2.09 ~ 2.35	30

Fig. 2 Scatter plot of spatial distribution for NDVI vs. TB, Band-Red, and Band-NIR



the data obtained, as Table 3 demonstrates. Moreover, the results of the correlation analysis reveal an improved correlation between TB and Red, as well as a correlation between TB and NDVI (Table 4), for which the correlation coefficients (CCs) are 0.61 and -0.439 , respectively. The NDVI is found to correlate negatively with TB, thus suggesting a large change in TB concentration as NDVI approaches 0.

Prediction Model

In this research, we use MLR analysis to develop a TB prediction model between images and in specific areas, as shown in Equations 5 and 6. Including NDVI was found to be able to increase the model’s R^2 to 0.667 and its explanatory power by 11.2 %, while reducing the RMSE of both the training examples and test examples to 47.125 (NTU) and 22.151 (NTU), respectively. Furthermore, the improvement rate was increased to 8.72 %, as shown in Table 5. Figure 3 shows that the model can predict unreasonable phenomena with negative TB because the TB concentration ranges from 1.23 to 499.36 NTU, which is a large change that would generally be prone to the inferior evaluation ability of the model for low TB concentrations. However, integrating NDVI can decrease errors in the test examples in order to improve the accuracy of the

prediction model, as shown in Fig. 4b, so that it can be applied as an index criterion for monitoring TB changes in reservoirs.

Sensitivity Analysis

Sensitivity analysis is a method that alters variables in a model within a specific range to study subsequent changes in the model. In this study, one specific variable is changed while other variables are substituted by mean values for the model in the research. By doing so, the change situation of the aforementioned specific variable in the model can be used to comprehend the sensitivity of various variables in the model. As shown in the sensitivity analysis diagram in Fig. 3, prior to integrating NDVI, the TB concentration change range of Band-Red is greater than that of Band-NIR, thus suggesting that Band-Red is more sensitive in such a prediction model; following NDVI’s integration, the TB concentration change range of Band-Red is still rather large, even though NDVI is playing a secondary role, as shown in Fig. 3b, thus indicating that NDVI is a sensitive factor in that prediction model as well.

Discussion and Conclusions

TB is a parameter of water quality. In previous studies, satellite bands are generally made to be variables in the model and are directly applied to determine TB

Table 3 Statistical table of image bands and in situ TB’s data

Statistics	Red (μm)	NIR (μm)	NDVI	TB (NTU)
MAX	62	35	-0.05	499.36
MIN	23	10	-0.46	1.23
AVG	34.7	21.44	-0.24	21.17
STD	8.6	6.2	0.1	75.21

Table 4 Correlation table of image bands and in situ TB’s data

CC	Red	NIR	NDVI
TB	0.61	0.049	-0.439

Table 5 Analysis of TB's prediction model

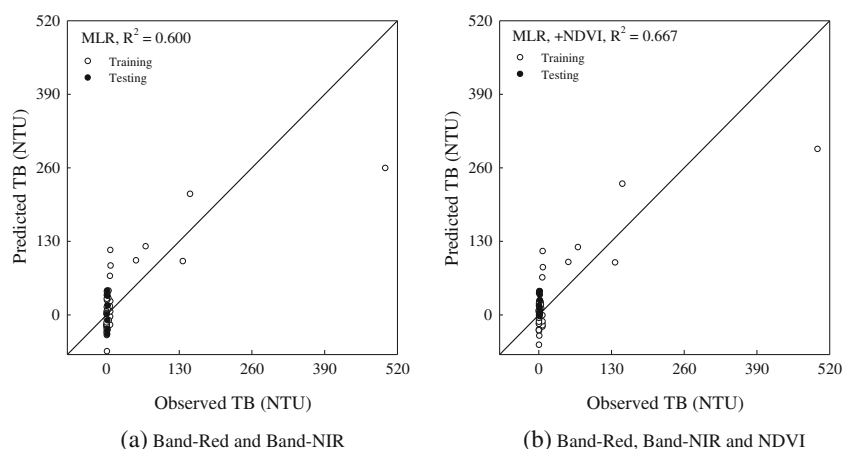
Statistics	TB	TB (+NDVI)
Training CC	0.775	0.817
Training R^2	0.6	0.667
Training RMSE (NTU)	51.63	47.125
Testing RMSE (NTU)	27.941	22.151

concentration. However, we found that NDVI, whose ratio type of Near Infrared Red (NIR) to Red falls between -1 and 1 , has a value closer to -1 in clean water and closer to 0 in turbid water. Consequently, only Red and NIR bands are used as investigative factors in this study in order to evaluate the possibility of using NDVI as a method for monitoring the TB change in reservoirs. Including NDVI was found to be able to increase the R^2 of the model to 0.667 , increase the explanatory power of the model by 11.2% , and reduce the RMSE of both the training samples and the test samples to 47.125 (NTU) and 22.151 (NTU), respectively, as well as increasing the improvement rate to 8.72% . Therefore, we can use the NDVI index as a variable in the model to accurately estimate TB concentration.

TB is an important parameter for monitoring the water quality of reservoirs. Since northern Taiwan's reservoirs are sensitive to SS, their characteristics can be monitored over a large range using water quality RS technology to better comprehend the spatial and temporal changes of TB. To create a prediction model, multiple single-bands, as well as NDVI, are used as investigative factors within this research. Regarding classification, the NDVI value is negative for water, while water quality has a negative correlation with TB. Although we only used Landsat-7 ETM+ satellite images here for the analysis data, said relationship is also found in data

analysis through SPOT-4 satellite images in later investigations. This result demonstrates that NDVI values approach 0 as TB increases, which indicates that NDVI can provide valuable information regarding TB changes, thus increasing the accuracy of the prediction model. However, only MLR analysis is used in this research to study the viability of integrating NDVI inclusion.

We believe that NDVI can be used as a variable to accurately estimate TB concentration, but MLR is a linear approach, and thus not being able to carry out nonlinear problems is among the methodological limitations of this research. With regard to future research, artificial intelligence (AI) tools, such as artificial neural networks (ANNs) and genetic algorithm operation trees (GAOTs), can be used in this field. ANN is an information processing system whose architecture mimics the brain's biological system; it is a relatively new computational tool that is especially useful when assessing systems with a considerable amount of nonlinear variables. On the other hand, GAOT, which consists of genetic algorithms (GA) and operation trees (OT), finds the best function and can explore complex relationships between inputs and outputs in the case that a physical model cannot be created in advance. Furthermore, AI tools may be included to not only enhance the monitoring ability of water quality by RS, but also assess the option of using NDVI to more completely monitor TB changes in reservoirs. Therefore, this model (with the NDVI variable) can be used to monitor a wide range of turbidity changes in a reservoir with RS. When TB concentration reveals abnormal changes in a reservoir, prevention efforts can be made in advance of a climate change. Furthermore, in the future, the model can be applied to determine TB changes in new satellite images and confirm our model's accuracy with regard to monitoring TB concentrations in reservoirs.

Fig. 3 Scatter figure of predicted TB's concentration

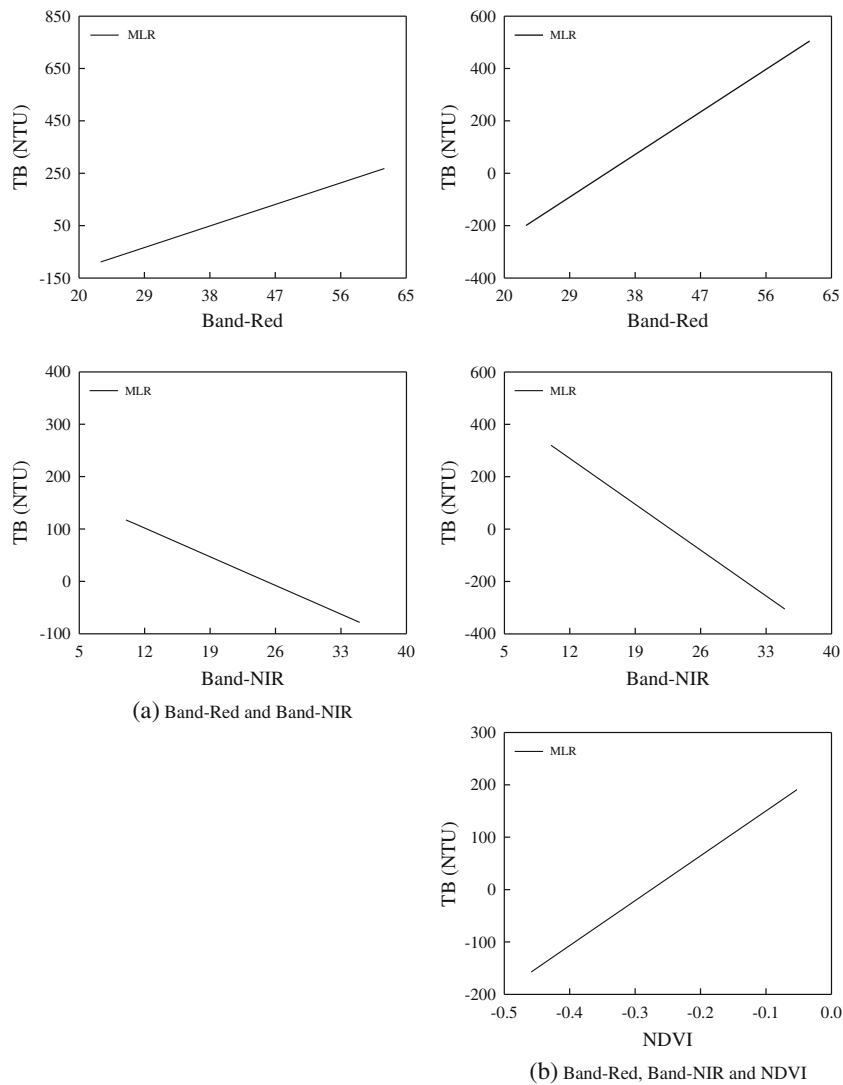


Fig. 4 The sensitivity figure of variables in TB’s prediction model

References

Chang, T. Y., Wang, Y. C., Feng, C. C., Ziegler, A. D., Giambelluca, T. W., & Liou, Y. A. (2012). Estimation of root zone soil moisture using apparent thermal inertia with MODIS imagery over a Tropical Catchment in Northern Thailand. *IEEE Trans Geoscience Remote Sensing*, 5, 752–761.

Changliang, Y. (2013). A weighted algorithm for estimating chlorophyll-a concentration from turbid waters. *Journal of the Indian Society of Remote Sensing*, 41(4), 957–967.

Dhillon, J. K., & Mishra, A. K. (2013). Estimation of trophic state index of Sukhna Lake using remote sensing and GIS. *Journal of the Indian Society of Remote Sensing*, 42(2), 469–474.

Dogliotti, A. I., Ruddick, K. G., Nechad, B., Doxaran, D., & Knaeps, E. (2015). A single algorithm to retrieve turbidity from remotely-sensed data in all coastal and estuarine waters. *Remote Sensing of Environment*, 156, 157–168.

Garaba, S.P., Badewien, T.H., Braun, A., Schulz, A.-C. & Zielinski, O. (2014) Using ocean colour remote sensing products to

estimate turbidity at the Wadden Sea time series station Spiekeroog. *Journal of the European Optical Society-Rapid Publications*, 9, 1–6, 14020.

He, W. Q., Chen, S., Liu, X. H., & Chen, J. N. (2008). Water quality monitoring in slightly-polluted inland water body through remote sensing - case study of the Guanting Reservoir in Beijing, China. *Environmental Science & Engineering in China*, 2(2), 163–171.

Koponen, S., Pulliainen, J., Kallio, K., Vepsalainen, J., & Hallikainen, M. (2001). Use of MODIS data for monitoring turbidity in Finnish Lakes. *IEEE*, 5, 2184–2186.

Liou, Y.-A., Sha, H.-C., Chen, T.-M., Wang, T.-S., Li, Y.-T., Lai, Y.-C., & Chiang, M.-H. (2012). Assessment of disaster losses in rice paddy field and yield after tsunami induced by the 2011 great East Japan Earthquake. *Journal of Marine Science and Technology (JMST)*, 20(6), 618–623.

Liou, Y.-A., Wang, T.-S. & Chan, H.-P. (2013). Impacts of pond change on the regional sustainability of water resources in Taoyuan, Taiwan, *Advances in Meteorology*, Article ID 243456, 6. doi:10.1155/2013/243456.

- Liou, Y.-A., Liu, H.-L., Wang, T.-S. & Chou, C.-H. (2015). Vanishing irrigation ponds and regional water resources in Taoyuan, Taiwan. *Terrestrial, Atmospheric and Oceanic Sciences (TAO)*, 26(2), Part II, 161–168.
- The official website of U.S. Geological Survey (USGS) Landsat project SLC-off products background. (2007). http://landsat7.usgs.gov/data_products/slc_off_data_products/slc_off_background.php.
- Wang, T.S., (2010). Application of various-scale satellite images to inland water quality monitoring in Taiwan. Ph.D. Thesis, Department of Civil Engineering and Engineering Informatics, Chung Hua University, 292 pp. (in Chinese).
- Wu, G., Leeuw, J. D., Skidmore, A. K., Prins, H. H. T., & Liu, Y. (2007). Concurrent monitoring of vessels and water turbidity enhances the strength of evidence in remotely sensed dredging impact assessment. *Water Research*, 41, 3271–3280.
- Zhang, Y., Pulliainen, J. T., Koponen, S. S., & Hallikainen, M. T. (2003). Water quality retrievals from combined Landsat TM data and ERS-2 SAR data in the Gulf of Finland. *IEEE*, 41(3), 622–629.