Development of New Water Quality Algorithms using Hyperspectral Data in Optically Complex Waters

光学特性が複雑な水域における超多波長反射率データを使った新しい水質推定手法の構築

September 2018

Zuomin Wang (王　作敏)

Contents

List of Figures

List of Tables

Abstract

Water is closely relative to our lives, to realize the monitoring and management of water quality for both inland and sea water is of great significance. Water quality parameters such as Chlorophyll-*a* (Chl-*a*), Total suspended solids (TSS) are important indicators to assess water quality, and sea surface salinity (SSS) is one of the key variables for monitoring and modeling ocean circulation. A variety of researches have been carried out to develop methods for detecting these water quality parameters using remote sensing methods in optically complex waters. However, there is not a unified algorithm applied for all kinds of waters due to the characteristic of regional dependency.

The objectives of this study are: (1) to develop models to estimate Chl-*a* and TSS in irrigation ponds in Higashihiroshima using *in situ* hyperspectral reflectance data and several regression analyses including (a) a simple linear regression at each waveband of reflectance and the first derivative reflectance (FDR); (b) all available two-band combination spectral indices (RSI and NDSI); and (c) a PLS regression with full spectra and ISE selected spectra (ISE-PLS) using original reflectance and FDR datasets (Chapter 3); (2) to evaluate the ISE-PLS method's accuracy to confirm the potential application to estimate Chl-*a* in the Seto Inland Sea comparing with several traditional approaches using *in situ* hyperspectral reflectance data (Chapter 4); (3) to test the applicability of $a_y(440)$ as a proxy for SSS in the bio-optical model, further, to establish the model for estimating SSS using *in situ* reflectance and water quality datasets in the Seto Inland Sea (Chapter 5); (4) to evaluate the bio-optical model developed with the Second generation Global Imager (SGLI) bands for SSS estimation in the Seto Inland Sea using *in situ* measurement data (Chapter6).

In Chapter 3, the author established quantitative models for estimating the Chl-*a* and the TSS concentrations in irrigation ponds in Higashihiroshima, Japan, using field hyperspectral measurements and statistical analysis. Field experiments were conducted in six ponds and spectral readings for Chl-*a* and TSS were obtained from six field observations in 2014. For statistical approaches, two spectral indices were used, the RSI and the NDSI, and a PLS regression. The predictive abilities were compared using the coefficient of determination (R^2), the root mean squared error of cross validation (RMSECV) and the residual predictive deviation (RPD). Overall, ISE–PLS, using FDR, showed the best predictive accuracy for both Chl-*a* ($R^2 = 0.98$, RMSECV = 6.15, RPD = 7.44) and TSS ($R^2 = 0.97$, RMSECV = 1.91, RPD = 6.64). The important wavebands for estimating Chl-*a* (16.97% of all wavebands) and TSS (8.38% of all wavebands) were selected by ISE–PLS from all 501 wavebands over the 400–900 nm range. These findings suggest that ISE–PLS based on field hyperspectral measurements can be used to estimate water Chl-*a* and TSS concentrations in irrigation ponds.

Harmful algal blooms (HABs) occur frequently in the Seto Inland Sea, bringing significant economic and environmental losses for the area, which is well known as one of the world's most productive fisheries. In Chapter 4, the author developed a quantitative model using *in situ* hyperspectral measurements in the Seto Inland Sea to estimate Chl-*a* concentration, which is a significant parameter for detecting HABs. Spectra and Chl-*a* data were obtained at six stations from 12 ship-based surveys between December 2015 and September 2017. In this study, the ISE-PLS regression method along with several empirical and semi-analytical methods such as ocean chlorophyll (OC), three-band model, and two-band model algorithms were used to retrieve Chl-*a*. Results showed that ISE-PLS using both the water-leaving reflectance (R_L) and FDR had a better predictive ability with higher R^2, lower root mean squared error (RMSE), and higher RPD values (R^2 = 0.77, RMSE = 1.47 and RPD = 2.1 for R_L; R^2 = 0.78, RMSE = 1.45 and RPD = 2.13 for FDR). However, in this study the OC algorithms had no predictive ability and the three-band and two-band model algorithms did not perform well in areas with lower Chl-*a* concentrations. These results support ISE-PLS as a potential coastal water quality assessment method using hyperspectral measurements.

SSS is playing a critical role for studying biological and physical processes in the ocean. In Chapter 5, the author established a model for estimating SSS using *in situ* measurement datasets from FY2015 to FY2016 in the Seto Inland Sea. To estimate SSS, a bio-optical model combined with the CDOM absorption coefficient at 440 nm ($a_y(440)$) was performed. Results showed that the estimated $a_y(440)$ had a reversely linear relationship with *in situ* SSS using winter datasets (R^2=0.66), which indicated the $a_y(440)$ is potential for modeling to estimate SSS, furthermore, the bio-optical model can be used for water quality assessment in Seto Inland Sea.

In Chapter 6, the author evaluated the bio-optical model combined with GCOM-C/SGLI bands to estimate SSS in the Seto Inland Sea. The *in situ* spectral and SSS data for the year of 2015 were analyzed using a two-band ratio and a bio-optical model. In the bio-optical model, the absorption of Gelbstoff in 380 nm (g_{380}) was utilized to build the relationship with SSS. According to the spectral pattern, the dataset were separated to two kinds of types (i.e., Type A and Type B). Results showed the estimated Chl-*a* had a linear relationship with observed Chl-*a* both for Type A and Type B dataset (R^2 = 0.69 for Type A and R^2 = 0.76 for Type B), and the estimated g_{380} also showed a linear relationship with observed SSS for both type of dataset (R^2 = 0.43 for Type A and R^2 = 0.64 for Type B), which showed the proposed bio-optical model based on SGLI bands has a potential predictive ability for Chl-*a* and SSS estimation in the Seto Inland Sea.

Chapter 1: General introduction

1.1. Background

Water quality is one of the main challenges that societies are facing at present, threatening human health, limiting food production, reducing ecosystem functions, and hindering economic growth (UNESCO, 2015). Water quality deterioration reflected in both inland and sea waters, which has become serious issue worldwide during the 21st century.

Inland waters including lakes and ponds are often subjected to sudden environmental changes caused by various anthropogenic reasons, such as agricultural, industrial, water supply, recreational, and touristic activities (El-Serehy et al., 2018). Agriculture is the greatest water consumer in the world, agricultural activities may bring pollution to the lakes and ponds, especially when using excess pesticides and fertilizers (FAO, 2009). There are about 21,000 irrigation ponds in Hiroshima Prefecture, which is the second largest number in Japan. In addition, a quarter of the total irrigation ponds in Japan are in Higashihiroshima (Abe & Shinohara, 1996), which indicates the irrigation ponds play an important role in agriculture for this area. It is essential to prevent and control deterioration of water quality to benefit ecological and human health and economic development (UNESCO, 2015).

Enclosed bays and coastal seas primarily face human occupied areas, anthropogenic activities directly influence the organisms living there (Irizuki et al., 2018). The Seto Inland Sea is the largest inland sea in Japan and one of the world’s productive fishery area. With the intensive industrializing and population density increasing around this area, severe deterioration in the water quality occurred in 1970s, particularly in seas facing highly populated and industrial area. As one of the famous marine disasters, red tides frequently occurred in the Seto Inland Sea during a period of high economic growth. Although red tides have decreased from about 300 cases per year in 1976 to about 100 cases per year more recently, severe damage to fisheries and significant economic losses are still occurring (Fukuyo et al., 2002).

Evaluating water quality can be defined as measurements of the physical, chemical, biological, and microbiological characteristics of water, which is used to alert us to current, ongoing, and emerging problems. Water quality can be evaluated according to indicators, to know well about water state information can help government or water manager making a sound decision to protect and manage water environment. Chlorophyll-*a* (Chl-*a*) is a key biochemical component in the molecular apparatus that is responsible for photosynthesis, presenting in many organisms including algae and some species of bacteria (WHO, 2011). It is widely used in ocean and inland waters as a proxy for phytoplankton biomass and indicator for eutrophication (Lu et al., 2016). Total suspended solids (TSS) is another one of the most commonly used indicators of surface water quality (Sikorska et al., 2015). It plays a fundamental role in inland waters as different materials including contaminants and pollutants can aggregate to these solids and brought in suspension (Giardino et al., 2017). Salinity has a direct effect on the stratification and circulation of the global ocean (Yu, 2011). As an indicator of seawater density, sea surface salinity (SSS) can help make better understanding in many physical and biogeochemical processes in coastal waters.

Due to the importance of irrigation ponds in Higashihiroshima and the Seto Inland Sea, to realize the assessment and monitoring of water quality is essential, especially a shorter temporal interval and wider range observation for water is needed. However, traditional measurement of water quality is time consuming and costly based on ship survey and laboratory analysis or a buoy system, in addition, sampling is often not enough to cover entire water body. Fortunately, remote sensing is generally more comprehensive than those directly measured *in situ* in that they provide greater spatial coverage with finer resolution and often increased temporal frequency and resolution (Kim et al., 2017).

Over the past two decades, various methods are applied in the estimation of Chl-*a* basing on remote sensed dataset, in both inland and coastal waters. In particular, satellites data basing on multiple spectral sensors are widely used, algorithms are developed according to various combinations of different spectral wavebands. As one of the most widely used methods, bio-optical model has been developed and successfully applied to satellite data to provide estimates of Chl-*a* concentration in coastal and oceanic waters (Shanmugam et al., 2018), which are based on the inherent optical properties (IOPs) in water. The ocean chlorophyll (OC) algorithms are widely used in open oceans as standard product algorithms for many satellite sensors, such as Sea-viewing Wide Field-of-view Sensor (SeaWiFS), and Moderate Resolution Imaging Spectroradiometer (MODIS). In addition, a reflectance ratio model using blue and green bands is acceptable to address the open oceans where phytoplankton drive the variation of IOPs of water, while inadequate for inland and coastal waters due to the influences from terrestrially materials (Zheng & DiGiacomo, 2017). Instead, another two-band model using near infrared (NIR) and red is applied in inland and coastal waters for Chl-*a* retrieval. A three-band model also has been proved available for inland water. TSS is another one of the most successful parameters that can be measured by means of remote sensing due to the effect of TSS on backscattering and water leaving radiance. As a consequent, a variety of empirical or semi-empirical methods have been developed using reflectance at appropriate wavebands as correlates. SSS are usually measured from microwave remote sensing in many past studies, particularly using the Soil Moisture and Ocean Salinity (SMOS) satellite. However, the resolution is too coarse (e.g. of the order of about 150 km of SMOS) to monitor SSS in coastal and estuarine zones (Qing et al., 2013).

Hyperspectral remote sensing has been used widely in recent years for its continuous waveband being very informative of water chemical feature. A great deal of researches are carried out using hyperspectral remote sensing method to retrieve various ground features. Partial least square (PLS) regression is a standard multivariate regression method developed by Herman Wold (1966), it has developed to become a standard tool in chemometrics and used in chemistry and engineering. PLS regression is particularly useful to handle a large number of descriptors even in the presence of co-linearity and noise (D'Archivioa et al., 2014). The basic principle is relating two data matrices, the predictor variables and the response variables, by a linear multivariate model, the original predictor variables are projected onto a small number of latent variables to simplify the relationship between both sides. Furthermore, the iterative stepwise elimination (ISE) method can be applied to remove useless or less useful predictor variables in PLS modelling to improve predictive performance. The combination of iterative stepwise elimination method and PLS (ISE-PLS) has proven to be an effective approach in chemistry. However, the potential application for water quality quantification has not been tested.

1.2. Objectives

A variety of researches have been carried out to develop methods for monitoring water quality using remote sensing methods in optically complex waters. However, there is not a unified algorithm applied for all kinds of waters due to the characteristic of regional dependency. In this study, the author aims to develop robust quantitative models using hyperspectral data to estimate water quality parameters in irrigation ponds in Higashihiroshima and the Seto Inland Sea.

In order to achieve the above goal, several important works have been conducted. Each chapter investigates about a particular issue and tries to solve its own specific objectives. The study consists of following objectives:

(1) To develop models to estimate Chl-*a* and TSS in irrigation ponds in Higashihiroshima using *in situ* hyperspectral reflectance data and several regression analyses including (a) a simple linear regression at each waveband of reflectance and the first derivative reflectance (FDR); (b) all available two-band combination spectral indices (RSI and NDSI); and (c) a PLS regression with full spectra and ISE selected spectra (ISE-PLS) using original reflectance and FDR datasets.

(2). To evaluate the ISE-PLS method's accuracy to confirm the potential application to estimate Chl-*a* in the Seto Inland Sea comparing with several traditional approaches using *in situ* hyperspectral reflectance data.

(3). To test the applicability of ay(440) as a proxy for SSS in the bio-optical model, further, to establish the model for estimating SSS using *in situ* reflectance and water quality datasets in the Seto Inland Sea.

(4). To evaluate the bio-optical model developed with SGLI bands for SSS estimation in the Seto Inland Sea using *in situ* measurement data.

1.3. Overview of the study areas

1.3.1. Irrigation ponds in Higashi-Hiroshima

The study area is located in Higashihiroshima, Japan, as shown in Fig. 1.1. Higashihiroshima is a core city in the central region of Hiroshima Prefecture, with a total area of 635.32 km^2 covering nearly 7.5% of the prefecture’s total area. Paddy fields, totalling 36.8 km^2, cover 14.9% of the Hiroshima Prefecture. Consequently, Higashihiroshima has the largest rice production of the 86 cities, towns and villages in Hiroshima Prefecture (Derbalah et al., 2003). The city has an estimated population of 183,834 people, and its population density was 289.36 people per km^2 in 2011. The number of irrigation ponds in Hiroshima Prefecture approaches approximately 21,000. This qualifies as the second largest number in Japan; a quarter of the total irrigation ponds in Japan are in Higashihiroshima, the average beneficiary area is 3.36 ha, and the average number of beneficiary farmhouses is approximately 9 (Abe & Shinohara, 1996). The monthly mean temperature ranges from 2.2 °C in January to 25.8 °C in August, and the monthly precipitation ranges from 43.3 mm in December to 232.1 mm in July, referring to the minimum and maximum values, respectively. To assess changes in water quality status and environments, six ponds, including both eutrophic ponds and non-eutrophic ponds, were selected for this study. Descriptions of the six ponds are listed in Table 1.1.

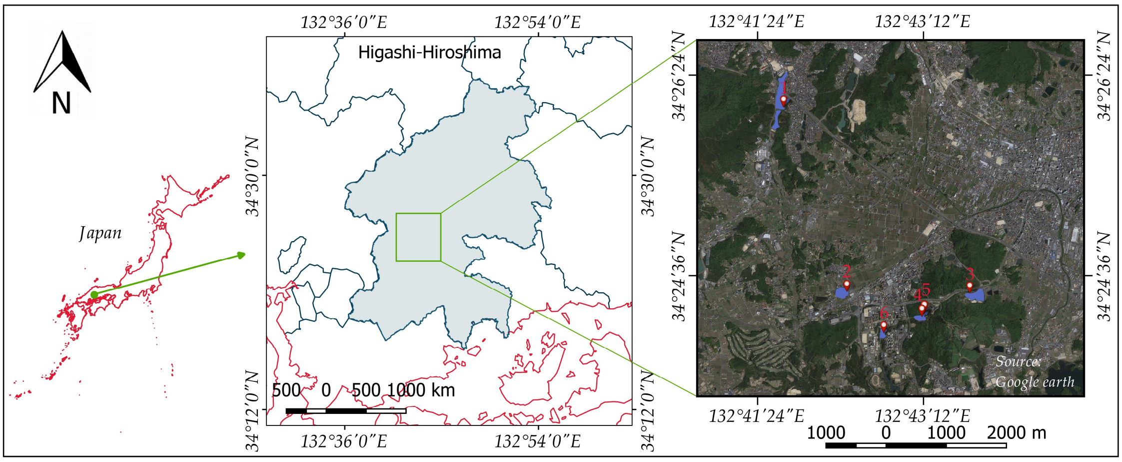

Figure 1.1. Locations of sampling sites in the irrigation ponds in Higashihiroshima, Japan

Table 1.1. Locations and brief information of sampling sites in the irrigation ponds in Higashihiroshima, Japan.

No.	Name of pond	Alt. (m)	Depth (m)	Area (ha)	Coordinate
1	Nanatsu-ike	245	2.3	8.1	34°26'06"N 132°41'39" E
2	Shitami-Oike	221	1.5	2.5	34°24'28"N 132°42'22"E
3	Okuda-Oike	228	3.3	2.9	34°24'25"N 132°43'43"E
4	Yamanaka-ike	231	2.6	1.2	34°24'14"N 132°43'12"E
5	Yamanakaike-kamiike	231	1.1	0.1	34°24'15"N 132°43'14"E
6	Budou-ike	210	1.6	1	34°24'02"N 132°42'45"E

1.3.2. The Seto Inland Sea

The Seto Inland Sea is the largest inland sea in Japan and one of the world's prominent inland seas (Fig. 1.2). It is surrounded by the three main Japanese islands: Honshu, Shikoku, and Kyushu, and contains approximately 700 small islands. The Seto Inland Sea had beautiful scenery with good water quality and high biodiversity; therefore, in 1934, it became the first designated national park in Japan. However, severe deterioration in the water and bottom qualities occurred, particularly in seas facing highly populated cities, and intensive industrialization in the coastal plains progressed from 1955 to 1973. The Seto Inland Sea is rich in fishery resources, with more than 50% of the total fish production contributed by aquaculture production. Additionally, there are approximately 17 fish farms being operated in the Tashima and Yokota areas near study area (Pawar et al., 2001). Consequently, there are a range of organic compounds released by fish farm waste that may affect eutrophication due to dissolved nitrogen (Pawar et al., 2002). The study area has an average water depth of 17.3 m and water temperatures range from 7.3°C (winter) to 28.4°C (summer). We selected six sites as sampling stations, which are described in Table 1.2.

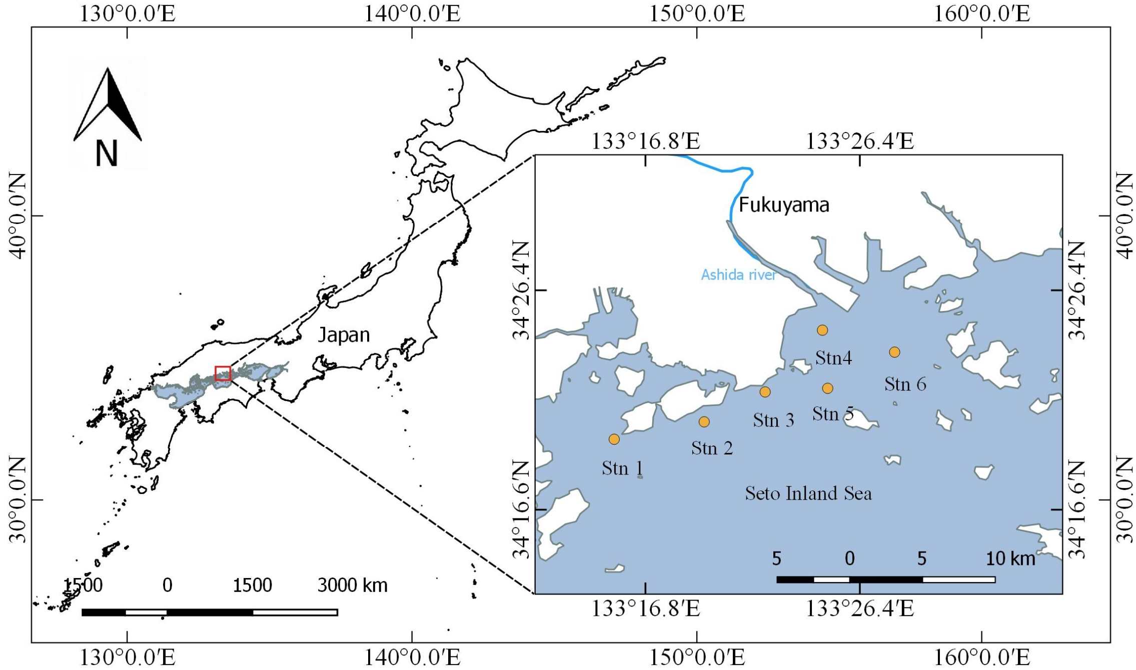

Figure 1.2. Locations of sampling stations in the Seto Inland Sea

Table 1.2. Locations and depths of sampling stations in the Seto Inland Sea.

Station ID	Latitude	Longitude	Depth(m)
1	34°19' 44" N	133°15' 24" E	29
2	34°20' 31" N	133°19' 26" E	17
3	34°21' 51" N	133°22' 10" E	10
4	34°24' 37" N	133°24' 44" E	10
5	34°22' 01" N	133°24' 58" E	21
6	34°23' 38" N	133°27' 58" E	17

1.4. Outline of the dissertation

The dissertation is composed of seven chapters. Chapter 1 clarifies the background, the objectives and the study areas about the present study. In chapter 2, several frequently used remote sensing algorithms and newly introduced algorithms for water quality assessment are explained, including the algorithms developed for detecting Chl-*a* and SSS. Chapter 3 describes a best fitted Chl-*a* and TSS retrieval method in irrigation ponds in Higashihiroshima. In Chapter 4, ISE-PLS method is evaluated in the Seto Inland Sea comparing with several frequently used methods. In chapter 5, an improved bio-optical method for SSS estimation in the Seto Inland Sea is developed. In chapter 6, a bio-optical model based on SGLI bands for SSS estimation is evaluated. In chapter 7, conclusions is described.

Chapter 2: Remote sensing algorithms for water quality

2.1. Traditional Chl-a algorithms

2.1.1. OC algorithm

The ocean chlorophyll algorithms (i.e., OC2, OC3, OC4) are based on the relationship between the Chl-a concentration and a blue to green band ratio of remote sensing reflectance R_{rs}. The band ratio of each algorithm consists of 2 to 4 bands in the spectrum region of 443 to 555 nm depending on the band element of the satellite sensor. Due to the subsurface reflectance in 555 nm is approximately independent of Chl-a concentration, the reflectance used in the empirical algorithms are usually expressed as ratios relative to their value at 555 nm, i.e., $R_{rs}(\lambda)/R_{rs}(555)$, λ is the wavelength. O'Reilly et al. (1998) developed the OC2 and OC4 algorithms for Sea-Viewing Wide Field-of-View Sensor (SeaWiFS) using a modified cubic polynomial relationship between Chl-a and a ratio of R_{rs}, which was the first version of OC algorithms. The newest OC algorithms (version 6) (Hu et al., 2012), yielded better statistical agreement between model data and Chl-a than the first version OC algorithm (O'Reilly et al., 1998), is formulated as a fourth-order polynomial with five coefficients and can be expressed as:

$$log_{10}(Chl\_a) = a_0 + a_1R + a_2R^2 + a_3R^3 + a_4R^4 \qquad (2.1)$$

where R is the ratio of optimal R_{rs} according to different sensors, the coefficients a_0, a_1, a_2, a_3, and a_4 are different in the OC2, OC3, and OC4 algorithms. Table 2.1 shows used bands and coefficients for several commonly used sensors (NASA, n.d.).

Table 2.1. Wavebands and coefficients used in OC algorithms for different sensors.

Algorithm	**Sensor**	**Blue band**	**Green band**	a_0	a_1	a_2	a_3	a_4
OC4	SeaWiFS	443>490>510	555	0.3272	-2.994	2.7218	-1.2259	-0.5683
OC4E	MERIS	443>490>510	560	0.3255	-2.7677	2.4409	-1.1288	-0.499
OC3S	SeaWiFS	443>490	555	0.2515	-2.3798	1.5823	-0.6372	-0.5692
OC3M	MODIS	443>488	547	0.2424	-2.7423	1.8017	0.0015	-1.228
OC3E	MERIS	443>490	560	0.2521	-2.2146	1.5193	-0.7702	-0.4291
OC3C	CZCS	443>520	550	0.333	-4.377	7.6267	-7.1457	1.6673
OC2S	SeaWiFS	490	555	0.2511	-2.0853	1.5035	-3.1747	0.3383
OC2M	MODIS	488	547	0.25	-2.4752	1.4061	-2.8233	0.5405
OC2E	MERIS	490	560	0.2389	-1.9369	1.7627	-3.0777	-0.1054

SeaWiFS, Sea-Viewing Wide Field-of-View Sensor; MERIS, MEdium Resolution Imaging Spectrometer; MODIS, Moderate Resolution Imaging Spectroradiometer; CZCS, Coastal Zone Color Scanner.

OC2 uses the blue/green ratio R_{rs} (blue)/R_{rs} (green) and R is described as $R = \log_{10}$ $(\frac{R_{rs}\text{ (blue)}}{R_{rs}\text{ (green)}})$. OC3 uses a three-band formulation with a maximum of R_{rs} band ratios R_{rs}(blue1)/R_{rs}(green) and R_{rs}(blue2)/R_{rs}(green), and R is expressed as $R = \log_{10}$ $(\frac{R_{rs}\text{ (blue1)>}R_{rs}\text{ (blue2)}}{R_{rs}\text{ (green)}})$. Similarly, OC4 uses the maximum of three R_{rs} ratios R_{rs}(blue1)/R_{rs}(green), R_{rs}(blue2)/R_{rs}(green) and R_{rs}(blue3)/R_{rs}(green)—to build the formulation, with R expressed as $R = \log_{10}$ $(\frac{R_{rs}\text{ (blue1)>}R_{rs}\text{ (blue2)>}R_{rs}\text{ (blue3)}}{R_{rs}\text{ (green)}})$.

2.1.2. Bio-optical model

The process of light absorption and scattering by the various seawater constituents such as particulate and dissolved organic matter are described by the inherent optical properties, IOPs, including absorption and backscattering coefficients, are the most significant parameters governing the light propagation within the water column and thus facilitate the estimation of aquatic biomass, primary production, heat flux, and carbon pools (Lee et al., 1996). In recent years, various models have been developed to relate IOPs to the apparent optical properties (AOPs), including irradiance reflectance and remote sensing reflectance R_{rs}, or to determine IOPs from AOPs (Gordon et al., 1975; Morel & Prieur, 1977; Gordon, 1991).

Remote sensing reflectance is a directional component to the light observed by the remote sensor, usually utilized to build relationship with IOPs, the relationship between remote sensing reflectance (just beneath the sea surface) and IOPs is defined as:

$$r_{rs}(\lambda) = \frac{L_u(\lambda, 0^-)}{E_d(\lambda, 0^-)} \tag{2.2}$$

where $L_u(\lambda, 0^-)$ is the upwelling or water leaving radiance, and $E_d(\lambda, 0^-)$ is the downwelling irradiance at null depth (denoted by 0^-). Gordon et al. (1988) described the relationship between r_{rs} and IOPs as:

$$r_{rs}(\lambda) = g_1\left(\frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)}\right) + g_2\left(\frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)}\right)^2 \tag{2.3}$$

where g_1 and g_2 are geometrical factors, $a(\lambda)$ and $b_b(\lambda)$ are total absorption and total backscattering coefficient, respectively (all symbols and abbreviations refer to Table 2.2). According to equation (2.3), a simplified equation is defined as

$$r_{rs}(\lambda) = \frac{f}{Q} * \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)} \tag{2.4}$$

where f is a factor of light field, Q is the light distribution factor defined as $Q = E_u/L_u$; E_u is the upwelling irradiance. The reflectance just above water, $R_{rs}(\lambda)$, usually related to $r_{rs}(\lambda)$ by an approximation of 0.54 (Austin, 1980), can be expressed as:

$$R_{rs}(\lambda) = 0.54 * \frac{f}{Q} * \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)} \tag{2.5}$$

The major light absorbing constituents in seawater are phytoplankton, detrital or nonalgal particles and colored dissolved organic matter (CDOM). The total absorption can be composed by the absorption of pure water a_w, phytoplankton absorption a_{ph}, nonalgal particles absorption a_{NAP}, and CDOM absorption a_{CDOM}, which can be expressed as:

$$a(\lambda) = a_w(\lambda) + a_{ph}(\lambda) + a_{NAP}(\lambda) + a_{CDOM}(\lambda) \tag{2.6}$$

where pure water absorption $a_w(\lambda)$ at each waveband is a specific value, which many researches have already measured (Pope & Fry, 1997). Phytoplankton absorption accounts for a significant proportion of the total absorption in oceanic waters and usually defined in terms of its pigment-specific coefficients as:

$$a_{ph}(\lambda) = [Chl\_a]\, a_{ph}^{*}(\lambda) \tag{2.7}$$

where $[Chl\_a]$ is Chl-a concentration, $a_{ph}^{*}(\lambda)$ is the specific absorption coefficient of phytoplankton. Many studies have demonstrated that the $a_{ph}^{*}(\lambda)$ values regularly decrease from 0.18 to 0.01 m^2/mg from oligotrophic to eutrophic waters which the chlorophyll concentration range 0.02 to 25 mg/m^3 (Bricaud et al., 1998). CDOM absorption has an exponential decay distribution with increasing wavelength and can be determined from their reference values at reference wavelength (e.g. 440 nm) as:

$$a_{CDOM}(\lambda) = a_{CDOM}(\lambda_0)e^{[-S_{CDOM}(\lambda-\lambda_0)]} \tag{2.8}$$

where $a_{CDOM}(\lambda_0)$ is the absorption coefficient for CDOM at a reference wavelength λ_0, S_{CDOM} is the slope coefficient of the exponential decrease and the value generally range from 0.01 to 0.02, vary with the CDOM concentration (Kirk, 1994). The absorption of nonalgal particles has proved to be similar to CDOM absorption as:

$$a_{NAP}(\lambda) = a_{NAP}(\lambda_0)e^{[-S_{NAP}(\lambda-\lambda_0)]} \tag{2.9}$$

where $a_{NAP}(\lambda_0)$ is the absorption coefficient of nonalgal particles, and S_{NAP} is the slope coefficient of the exponential decrease for nonalgal particles.

The total backscattering coefficient can be separated into contributions by backscattering of pure water $b_{bw}(\lambda)$ and particles $b_{bp}(\lambda)$, and can be expressed as:

$$b_b(\lambda) = b_{bw}(\lambda) + b_{bp}(\lambda) \tag{2.10}$$

where $b_{bw}(\lambda)$ is the backscattering due to water molecules, $b_{bp}(\lambda)$ is particle backscattering.

Table 2.2. Symbols and definitions

Symbol	Description	unit
$L_u(0^-)$	Upwelling radiance beneath water surface	W m^{-2} sr^{-1}
$E_d(0^-)$	Downwelling irradiance	W m^{-2}
$a(\lambda)$	Total absorption coefficients	m^{-1}
$a_w(\lambda)$	Absorption coefficients of pure water	m^{-1}
$a_{ph}(\lambda)$	Absorption coefficients of phytoplankton	m^{-1}
$a_{NAP}(\lambda)$	Absorption coefficients of nonalgal particles	m^{-1}
$a_{CDOM}(\lambda)$	Absorption coefficients of CDOM	m^{-1}
$b_b(\lambda)$	Total backscattering coefficients	m^{-1}
$b_{bw}(\lambda)$	Backscattering coefficients of pure water	m^{-1}
$b_{bp}(\lambda)$	Backscattering coefficients of particles	m^{-1}
$R_{rs}(\lambda)$	Remote sensing reflectance above water surface	sr^{-1}
$r_{rs}(\lambda)$	Remote sensing reflectance beneath water surface	sr^{-1}

2.1.3. Three-band model

The three-band model, developed by Gitelson et al. (2003) and initially applied in terrestrial vegetation, has been proved to be a useful method to assess Chl-a in turbid productive waters (Dall'Olmo et al., 2003). This method is based on the relationship between reflectance $R(\lambda)$ and two IOPs, i.e., total absorption and backscattering coefficients. The relationship can be expressed as:

$$R(\lambda) \propto \gamma \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)} \tag{2.11}$$

where γ is dependent on the geometry of the light field emerging from the water body. The relationship between Chl-a and three band model can be expressed as:

$$\mathrm{Chl}\_a \propto (R_{\lambda1}^{-1} - R_{\lambda2}^{-1}) \times R_{\lambda3} \tag{2.12}$$

where $R_{\lambda \mathrm{i}}$ is the reflectance at a wavelength of λi nm. We found the optimal spectral ranges for these wavelengths to be, λ1 = 660–670 nm, which is maximally sensitive to absorption by Chl-a for the reason of building the relationship with Chl-a absorption coefficient by the conceptual of $R(\lambda)$; λ2 = 690–720 nm, which is minimally sensitive to absorption by Chl-a and can be used to decrease the impact by absorption of CDOM and nonalgal particles; and λ3 = 720–750 nm, which is minimally affected by absorption by any constituents and used for decrease the impact by backscattering (Dall’Olmo & Gitelson, 2005; Gitelson et al., 2008).

2.1.4. NIR/Red ratio

The NIR/red two-band model has been widely used to retrieve Chl-a concentrations in turbid productive waters to identify phytoplankton blooms (Gitelson A. A., 1992). This model is a special case of the three-band model, when the Chl-a absorption coefficient is much larger than the backscattering coefficient and Chl-a absorption coefficient is much large than the sum of the absorption coefficient of CDOM and NAP. The model can be formulated as follows:

$$\mathrm{Chl}\_a \propto R_{\lambda1}^{-1} \times R_{\lambda2} \tag{2.13}$$

where λ_1 is in the red region and λ_2 is a NIR band which is similar with λ_3 in the three-band model.

2.2. Traditional Salinity algorithms

2.2.1. Microwave method

Ocean salinity has a significant impact on the density of seawater, and then density is one of the key influencing variables for ocean circulation, which plays a crucial role in moderating the climate. SSS is also critical for determining the global water balance and for estimating evaporation rates. In addition, salinity variations are dominantly driven by precipitation, evaporation, and runoff, ice freezing and melting is also an important factor in some higher latitude area. Consequently, SSS is a variable that not only can provide valuable estimations of rainfall over the oceans but also is fundamental in other processes that force our global climate system (Klemas, 2011).

However, there is a lack of *in situ* salinity time series measurements over most of the global ocean, and only a small fraction of the ocean is sampled on a regular basis. Recently, airborne based microwave radiometers have been applied to estimate SSS, for instance, airborne scanning low-frequency microwave radiometers (SLFMRs) and Salinity, Temperature, and Roughness Remote Scanner (STARRS). Both instruments are based on the physical basis for microwave radiometry of the ocean, which can be explained by an equation formulated as:

$$T_b = eT \tag{2.14}$$

where T_b is the ocean radiometric brightness temperature; T is thermodynamic temperature; and e represents the emissivity, which is a function of seawater conductivity. Because salinity is crucial for the conductivity, SSS can be estimated by retrieval of T_b and sea surface temperature (SST). The SLFMRs and STARRS are both use an L-band radiometer for the primary brightness temperature measurement and an infrared radiometer for SST.

Spacecraft remote sensing of SSS, for instance, Soil Moisture and Ocean Salinity (SMOS) satellite, launched by the European Space Agency in 2009, is also estimated by means of T_b using the Microwave Imaging Radiometer using Aperture Synthesis (MIRAS), which is a novel interferometric radiometer that operates in the L-band microwave range. The SMOS satellite maps salinity with 0.1 practical salinity units (psu), by averaging over 10–30 days in areas measuring 200 × 200 km. Although satellite-based microwave radiometry method can map SSS with a wide range, the spatial resolution is too coarse to monitor SSS in coastal and estuarine areas.

2.2.2. Optical method

Unlike the L-band space-borne radiometers, optical multispectral satellites have better spatial resolution. For instance, the spatial resolution of MEdium Resolution Imaging Spectrometer (MERIS) and MODIS range to several hundred meters, further, the Landsat satellite can reach to 30 m for each pixel. Several studies have been conducted for SSS retrieval using above satellite data, mainly all using a multiple linear relationship between multispectral bands and SSS (Khorram, 1982; Wong et al., 2007; Qing et al., 2013). SSS has been proved inverse linearly relates to CDOM concentration in coastal or estuarine waters (Bowers et al., 2000). Recently, SSS has been derived using the absorption coefficient of CDOM, basing on the relationship between the absorption coefficient of CDOM and R_{rs} derived from *in situ* measurement (Ahn et al., 2008).

2.3. Proposed Chl-*a* algorithm

PLS method was developed in the 1960's by Herman Wold, originated in the social sciences but became popular first in chemometrics. The traditional PLS regression equation is described as follows:

$$y_i = \beta_{i1}x_{i1} + \beta i_2 x_{i2} + \cdots + \beta_{im}x_{im} + \varepsilon = \sum_{j=1}^{m} \beta_{ij}x_{ij} + \varepsilon \qquad (2.15)$$

where y_i ($i = 1,2, \ldots, n$) is the response variable that represents Chl-*a*, and x_{ij} ($j = 1,2, \ldots, m$) is the predictor variable representing spectral data such as reflectance values for spectral bands 1 to m. β_n is the estimated weighted regression coefficient, and ε is the error vector. In the PLS model, the original predictor variables (X) are projected onto a small number of orthogonal latent variables to simplify their relationships with response variables (Y) (Li & He, 2008). The progress of PLS calculate the latent variable (t_{ik}) ($k = 1,2, \ldots, r$) shows below:

$$y_i = c_1t_{i1} + c_2t_{i2} + \cdots + c_mt_{ir} + \varepsilon = \sum_{k=1}^{r} c_kt_{ik} + \varepsilon \qquad (2.16)$$

where

$$t_{i1} = w_{11}x_{i1} + w_{12}x_{i2} + \cdots + w_{1m}x_{im} + \varepsilon = \sum_{j=1}^{m} w_{1j}x_{ij}$$

$$\vdots \qquad (2.17)$$

$$t_{ir} = w_{r1}x_{i1} + w_{r2}x_{i2} + \cdots + w_{rm}x_{im} + \varepsilon = \sum_{j=1}^{m} w_{rj}x_{ij}$$

The latent variable (t_{ik}) acts as a role of medium, then combine the equation (2.16) and (2.17):

$$
\begin{aligned}
y_i &= c_1(w_{11}x_{i1} + w_{12}x_{i2} + \cdots + w_{1m}x_{im}) + \cdots \\
&\quad + c_r(w_{r1}x_{i1} + w_{r2}x_{i2} + \cdots + w_{rm}x_{im}) + \varepsilon_i \\
&= \left(\sum_{k=1}^{r} c_k w_{k1}\right) x_{i1} + \left(\sum_{k=1}^{r} c_k w_{km}\right) x_{im} + \varepsilon_i \\
&= \sum_{j=1}^{m}\left(\sum_{k=1}^{r} c_k w_{kj}\right) x_{ij} + \varepsilon_i
\end{aligned}
\tag{2.18}
$$

where

$$\beta_j = \sum_{k=1}^{r} c_k w_{kj} \tag{2.19}$$

then, the relation between water quality parameters (y_i) and wavebands (x_i) is same with PLS regression equation (2.15)

We selected the optimal number of latent variables in the final model using the leave-one-out (LOO) cross-validation method with a minimum value of the root mean squared error (RMSE), which is calculated as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(y_i - y_p\right)^2}{n}} \tag{2.20}$$

where y_i and y_p represent sample i's measured and predicted Chl-*a*, respectively, and n is the number of samples in the dataset ($n = 59$).

The ISE-PLS uses a model-wise elimination technique (Boggia et al., 1997) that permits the removal of useless descriptors to improve predictive performance. This process is based on the importance of the predictor z_i, which is defined as:

$$z_i = \frac{|\beta_i| s_i}{\sum_{i=1}^{I} |\beta_i| s_i} \tag{2.21}$$

where s_i is the standard deviation of predictor x_i. PLS modeling uses all available wavebands (501 bands between 400 and 900 nm). Predictors are then evaluated based on the value of the importance of predictor z_i. The predictor with minimum importance is eliminated in each elimination cycle and the remaining predictors are used to recalibrate the model (Forina et al., 2003). Finally, a model with maximum predictive ability is selected using the minimum RMSE value from the cross-validation.

2.4. Proposed Salinity algorithm

Bio-optical model is widely used to predict the water quality parameters from the spectral observation. Water quality parameters (e.g. Chl-a, CDOM, SS) can be estimated when the relationship between IOPs and spectral reflectance is built. However, SSS can't be estimated directly using this method, because the salinity is not included in the total absorption coefficient in water, and ocean salinity is not easily observed using optical satellite for the reason that there is no single band which correlates highly with *in situ* salinity data (Wong et al., 2007). Fortunately, Bower et al. (2000) found SSS linearly relates to CDOM concentration in coastal or estuarine waters. This finding provide possibility for SSS estimation using the absorption coefficient of CDOM in bio-optical model.

A bio-optical model performed well for Chl-a and CDOM estimation in Tokyo bay (Sugihara et al., 1985). The total absorption contributed by pure water, Chl-a, and CDOM, and can be expressed as:

$$a(\lambda) = a_w(\lambda) + a_c(\lambda)[Chl_a] + a_y(\lambda_0)\exp\{-S(\lambda - \lambda_0)\} \tag{2.22}$$

where $a(\lambda)$ is the total absorption coefficient of the sea water, $a_w(\lambda)$ is the absorption coefficient of pure sea water, $a_c(\lambda)$ is the specific absorption coefficient of Chl-*a*, $a_y(\lambda_0)$ is the absorption coefficient of CDOM at wavelength λ_0. The backscattering can be formulated as:

$$b_b(\lambda) = b_w(\lambda)/2 + B_p \tag{2.23}$$

where $b_b(\lambda)$ is the total backscattering coefficient, $b_w(\lambda)$ is the scattering coefficient of pure sea water and B_p is the backscattering coefficient of particles.

$R_{rs}(\lambda)$ is remote sensing reflectance at the wavelength λ just above the water, which usually obtained by satellite sensor or other instruments is used to predict the water quality parameters. $R_{rs}(\lambda)$ has a relationship with remote sensing reflectance just beneath water surface $r_{rs}(\lambda)$ as:

$$R_{rs}(\lambda) = 0.52 * r_{rs}(\lambda)/(1 - 1.7 * r_{rs}(\lambda)) \tag{2.24}$$

Morel and Prieur (1977) described the relationship between spectral reflectance and IOPs and can be expressed as:

$$r_{rs}(\lambda) = 0.33 * b_b(\lambda)/a(\lambda) \tag{2.25}$$

Combining the above equations, we can get the relation as follows,

$$R_{rs}(\lambda) = \frac{0.33}{\pi} * \frac{b_w(\lambda)/2 + B_p}{a_w(\lambda) + a_c(\lambda)[Chl\_a] + a_y(\lambda_0)exp\{-S(\lambda - \lambda_0)\}} \tag{2.26}$$

if we introduce the X as the coefficients, the following equation can be obtained,

$$X_1[Chl\_a] + X_2[B_p] + X_3[a_y(440)] = Y \tag{2.27}$$

where

$X_1 = R_{rs}(\lambda)a_c(\lambda)/\pi$

$X_2 = -0.33/\pi$

$X_3 = R_{rs}(\lambda) exp\{-0.014(\lambda - 440)\}/\pi$

$Y = 0.33 * b_w(\lambda) 2\pi - a_w(\lambda) R_{rs}(\lambda)/\pi$

when we obtain corrected R_{rs} data beyond 3 bands, it will be possible to retrieve the three unknown parameters Chl-a, B_p and a_y440 at the same time to solve a simultaneous linear equation.

Chapter 3: Development of new chlorophyll-*a* algorithm in Ponds using ISE-PLS method

3.1. Introduction

Agriculture is by far the greatest water consumer in the world, and consequently, a major cause of water pollution. The primary pollutants from agriculture are excess nutrients and pesticides (Mateo-Sagasta & Burke, 2010). In agricultural activity, non-point source pollution, such as irrigation water and surface runoff water containing fertilizer from farmland, contributes to excessive nutrient concentrations (Yang et al., 2008). Meanwhile, excess nutrients that cause eutrophication, hypoxia and algal blooms in surface water bodies and coastal areas contribute to the primary global water quality problem (Mateo-Sagasta & Burke, 2010). Eutrophication has become a widespread matter of concern during the past 50 years, especially in coastal and inland waters (Rönnberg & Bonsdorff, 2004).

The Chl-*a* concentration in water is the most widely applied parameter to assess the water quality status of lakes, particularly with respect to their trophic quality (WHO, 2011). Since Chl-*a* is the primary photosynthetic pigment of all plant life (Latif et al., 2003), the concentration of Chl-*a* indicates phytoplankton biomass and eutrophication in lakes (Lu et al., 2016). The concentration of total suspended solids (TSS) is another commonly used indicator for water quality assessment (Sikorska et al., 2015). TSS consists of organic and inorganic materials suspended in the water (Fondriest Environment, Inc., 2016). Increased TSS decrease light transmission through the water

(Bash et al., 2001), and therefore affect light availability to phytoplankton, thus resulting in a decrease of phytoplankton primary production (Davies‐Colley & Smith, 2007).

However, traditional water quality monitoring requires *in situ* measurements and sampling, then returning the samples to the laboratory to measure water quality indicators (e.g., Chl-*a* and TSS), which is costly and time consuming (Shafique et al., 2003). Remote sensing makes it possible to monitor the state of the globe routinely, and is cost effective and useful, with the benefits of its passive nature and wide spatial coverage (Voutilainen et al., 2007). Earlier studies have demonstrated several algorithms developed for satellite sensors to estimate ocean and coastal water quality parameters, such as the Chl-*a* algorithm OC3, created for MODIS data, and OC4, created for sea-viewing wide field-of-view sensor (SeaWiFS) data (O'Reilly et al., 1998). The geostationary ocean color imager (GOCI) also shows good performance, using the linear combination index (LCI) method to monitor Chl-*a* (Sakuno et al., 2013). Further, a three-band semi-analytical reflectance model, originally developed by Gitelson et al. (2003), and a normalized difference chlorophyll index (NDCI) (Mishra & Mishra, 2012), both performed well for assessing Chl-*a* in turbid productive water (Mishra & Mishra, 2012; Dall'Olmo et al., 2003; Gitelson et al., 2008). For estimating TSS concentrations, an algorithm with a single wavelength created for MODIS and MERIS data has been proved to be satisfactory (Nechad et al., 2010).

Unlike ocean and coastal water, inland water usually has a smaller surface area and more complicated spectral features, especially irrigation ponds, which are often impacted by human use such as agriculture activities. Consequently, inland water quality monitoring presents higher requirements for both temporal and spatial resolution of satellite sensor data; hence currently used satellite sensors often have limited practical applicability in assessing relatively smaller inland water bodies. Since there are a limited number of wavebands for Landsat and other multispectral sensors, finding more informative wavebands to improve the performance of water quality estimation is necessary. With respect to *in situ* measurements, a two-band ratio approach, for example the ratio spectral index (RSI), has performed well for estimating Chl-*a* concentrations in inland waters (Gitelson et al., 2008; Nechad et al., 2010; Pulliainen et al., 2001), especially using the ratio of near-infrared (NIR) regions to red wavebands, such as the reflectance ratio of 705 nm to 670 nm performed by Han and Rundquist (1997). Normalized difference spectral indices (NDSI) are another type of spectral indices frequently used to select the optimum bands for spectral analysis. As similar studies that have been done before mainly focused on vegetation parameters retrieval (Inoue et al., 2008; Stagakis et al., 2010; Inoue et al., 2012), optimum bands have been calculated from combinations of all available bands in the hyperspectral spectrum, a considerable range for hyperspectral analysis. Water quality parameters retrieval requires a similarly broad approach.

Partial least squares (PLS) regression, which was developed by Wold (1966), is widely used to extract valuable information for spectroscopic analysis. PLS regression uses all available wavebands without multi-collinearity issues. The eigenvectors of the explanatory variables are manipulated such that the corresponding scores (latent variables) not only explain the variance of the explanatory variables (wavebands) themselves, but also are highly correlated with the response variables (Chl-*a* and TSS) (Song et al., 2014). However, PLS is considered limited because it treats each wavelength as independent, which incorporate noise created by non-informative wavelengths (Ghasemi et al., 2003). There is increasing evidence to indicate that wavelength selection can affect the performance of PLS analysis (Kawamura et al., 2008), since wavelength selection for PLS models is performed to eliminate uninformative variables and choose the variables that contribute the most to the predictive ability of the calibration model (Swierenga et al., 1998). Iterative stepwise elimination PLS (ISE–PLS), developed by Boggia et al. (1997), combines PLS regression and the most useful information from hundreds of wavebands into the first several factors (Kawamura et al., 2010; Derbalah et al., 2003; Abe & Shinohara, 1996). This method was developed to eliminate useless wavebands in PLS analysis.

The objective of this study is to develop models to estimate Chl-*a* and TSS using *in situ* spectral reflectance data and statistical approaches. We used several regression analyses including (a) a simple linear regression at each waveband of reflectance and the first derivative reflectance (FDR) to explore informative wavelength regions for Chl-*a* and TSS estimation; (b) all available two-band combination spectral indices (RSI and NDSI); and (c) a PLS regression using original reflectance and FDR datasets. In the PLS analyses, the predictive ability of ISE–PLS was compared with that of a standard full spectrum PLS (FS–PLS) and the spectral indices (RSI and NDSI).

3.2. Study area

The study area is located in Higashihiroshima, Japan, as shown in Fig. 1.1 in Chapter 1. To assess changes in water quality status and environments, six ponds, including both eutrophic ponds and non-eutrophic ponds, were selected for this study. Descriptions of the six ponds are listed in Table 1.1 in chapter 1.

3.3. Materials and methods

3.3.1. Measurement of Water Surface Reflectance

Measurements of water surface reflectance were performed using an ASD FieldSpec HandHeld-2 spectrometer (ASD Inc., Boulder, CO, USA) with a spectral range of 350–1050 nm and a probe field angle of 10°. Spectral readings were taken approximately 1 m above the water surface between 10:30 and 13:00 on a day with clear skies. Surveys were conducted six times between 3 January 2014, and 28 June 2014. From these data, a total of 36 datasets were obtained.

With respect to the spectral data, the ranges 325–399 nm and 901–1075 nm from each spectrum were identified as noise and removed. Subsequently, spectral data were smoothed using a moving and normalized Gaussian filter with a sigma (standard deviation) of 2.5. The FDR was also computed and compared with the original reflectance.

3.3.2. Water Sampling and Chemical Analysis

The water sampling sites were consistent with the spectral reflectance measurements. Immediately after measurement of spectral reflectance, water samples were collected into two 1 L containers. The samples were maintained at constant temperature and protected from light until they were received at the laboratory for analysis.

Chl-*a* and TSS concentrations were determined at the laboratory of the Graduate School for International Development and Cooperation (IDEC), Hiroshima University, Japan. Chl-*a* was extracted using 90% acetone, the absorption of Chl-*a* was measured by a spectrophotometer (UVmini-1240, SHIMADZU Co., Kyoto, Japan) and pigment concentration was calculated using the equations from UNESCO. To measure the TSS, the water sample was filtered using 47 mm diameter GF/F filters. The filters were weighed before and after drying with an oven drier (SANYO Electric Co., Moriguchi, Osaka, Japan) at 105 °C for two hours. The TSS contents were quantified by the difference in the weight of the filter paper before and after filtration.

3.3.3. Ratio Spectral Index and Normalized Difference Spectral Indices

A combination of spectral indices between all wavebands is performed to select the optimum two-band combination. The aim of spectral indices is to construct a mathematical combination of spectral wavebands to enhance information content with respect to the parameter under study (Stratoulias et al., 2015). Moreover, normalization in the NDSI is effective at cancelling atmospheric disturbance or other sources of error, while enhancing and standardizing the spectral response to the observed targets (Inoue et al., 2008).

For this study, two of the most commonly used spectral indices (RSI and NDSI) were calculated using the reflectance dataset. The forms to express them are as follows:

$$\mathrm{RSI}(i,j) = \frac{R_i}{R_j} \tag{3.1}$$

$$\mathrm{NDSI}(i,j) = \frac{R_i - R_j}{R_i + R_j} \tag{3.2}$$

3.3.4. Full Spectrum Partial Least Squares Regression

We performed PLS regression to estimate Chl-*a* and TSS concentrations using the reflectance and FDR datasets (n = 36). The standard FS–PLS regression equation is as follows:

$$y = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_i x_i + \varepsilon \tag{3.3}$$

where the response variable y is a vector of the water quality parameters (Chl-*a* and TSS), the predictor variables x1 to xi are surface reflectance or FDR values for spectral bands 1 to i (400, 401, …, 900 nm), respectively, β1 to βi are the estimated weighted regression coefficients, and ε is the error vector. The latent variables were introduced to simplify the relationship between response variables and predictor variables. To determine the optimal number of latent variables (NLV), leave-one-out (LOO) cross validation was performed to avoid overfitting of the model, which was based on the minimum value of the root mean squared error (RMSECV). The RMSECV is calculated as follows:

$$\text{RMSECV} = \sqrt{\frac{\sum_{i=1}^{n}\left(y_i - y_p\right)^2}{n}} \tag{3.4}$$

where yi and yp represent the measured and predicted water quality parameters (Chl-*a* and TSS) for sample i, and n is the number of samples in the dataset (n = 36).

3.3.5. Iterative Stepwise Elimination Partial Least Squares Regression

The ISE–PLS is a model-wise technique (Boggia et al., 1997), which is based on the wavelengths selection function of the ISE method. To improve the performance of the

PLS model, the optimum wavelengths with good predictive ability are selected for model calibration. The wavelengths elimination process depends on the importance of the predictors (zi), described as follows:

$$z_i = \frac{|\boldsymbol{\beta}_i|\boldsymbol{s}_i}{\sum_{i=1}^{I}|\boldsymbol{\beta}_i|\boldsymbol{s}_i} \tag{3.5}$$

where βi is the regression coefficient and si is the standard deviation of predictor, both corresponding to the predictor variable of the waveband i.

Initially, all available wavebands (501 bands, 400–900 nm) are used to develop the PLS regression model. Then variables are ranked from most contributed to least contributed according to the predictor zi; in other words, the predictor zi represents the weight of each variable. The least contributed variable is eliminated and the PLS model is recalibrated with the remaining predictor variables (Forina et al., 2004). The model building procedure is repeated, and in each cycle the predictor variable with the minimum importance (i.e., the less informative wavelength) is eliminated, until the final variable is eliminated. To determine the optimum number of wavelengths to include in the final model, LOO cross validation is conducted after each calibration. The final model with the maximum predictive ability is calibrated by the minimum value of RMSECV (D'Archivioa et al., 2014).

3.3.6. Evaluation of Predictive Ability

The coefficient of determination (R^2) and RMSECV were selected as indices to evaluate the FS–PLS and ISE–PLS calibration models' accuracy by using LOO cross validation. High results for R^2 and low RMSECV indicate the best model to predict Chl-*a* and TSS concentrations. In addition, the residual predictive deviation (RPD) was used to evaluate the predictive ability of the models, which was defined as the ratio of standard deviation (SD) of reference data in prediction to RMSECV (Williams, 2001). For determining the performance ability of the calibration models, the goal RPD was at least 3 for agriculture applications; RPD values between 2 and 3 indicate a model with good prediction ability, 1.5 < RPD < 2 is an intermediate model needing some improvement, and an RPD < 1.5 indicates that the model has poor prediction ability (D'Acqui et al., 2010).

All data handling and linear regression analyses were performed using Matlab software ver. 8.6 (MathWorks, Sherborn, MA, USA).

3.4. Results

3.4.1. Chl-*a* and TSS Concentrations in Irrigation Ponds

Descriptive statistics are shown in Table 3.1, including the sampling data, the number of samples, the minimum (Min), the maximum (Max), the mean, the standard deviation (SD) and the coefficient of variation (CV). In total, 36 samples were collected from six irrigation ponds in six sets of field measurements (3 January, 19 January, 24 March, 9 April, 24 May, and 28 June in 2014). Field samples (n = 36) provided a wide range of both Chl-*a* (SD = 46.1 μg/L, CV = 2.0) and TSS (SD = 12.8 mg/L, CV = 1.65). In the datasets, Chl-*a* ranged from 0 to 169.5 μg/L, and TSS ranged from 0.1 to 53 mg/L, which indicates that this study involves various water quality conditions from different ponds.

Table 3.1. Descriptive statistics for the Chl-*a* and TSS concentrations.

Date	***n***	**Chl-*a* (µg/L)**					**TSS (mg/L)**				
		Min	**Max**	**Mean**	**SD**	**CV**	**Min**	**Max**	**Mean**	**SD**	**CV**
3-Jan-14	6	0.1	98.7	20.7	39.1	1.9	0.1	16.8	6.1	7.2	1.2
19-Jan-14	6	0.1	169.5	36	67.5	1.9	0.1	26.5	7.6	11	1.5
24-Mar-14	6	0	169.1	36.8	67.3	1.8	0.4	38	10.2	15.5	1.5
9-Apr-14	6	0.5	48.5	8.7	19.5	2.2	0.5	33.5	6.5	13.2	2
24-May-14	6	0.9	37.7	9.2	14.6	1.6	0.2	26	5.8	10	1.7
28-Jun-14	6	1.6	133.9	27.1	52.5	1.9	0.3	53	10.4	20.9	2
Total	36	0	169.5	23.1	46.1	2	0.1	53	7.8	12.8	1.7

SD = standard deviation; CV = coefficient of variation; *n* = number of samples.

3.4.2. Comparison of Simple Linear Regression Models

In this study, several simple linear regression models were constructed, and the accuracy was compared with that of the PLS method. As shown in Table 3.2, distinct bands were selected as the optimal bands with respect to accuracy for all models. In the model that used the gaussian smoothed water surface reflectance and FDR, the 730 nm and 705 nm wavebands were selected, based on the linear correlation coefficient shown in Fig. 3.1, to estimate Chl-*a* concentration (R^2 = 0.14 and 0.54); 722 nm and 704 nm were selected to estimate TSS (R^2 = 0.05 and 0.46). Fig. 3.1 shows the correlation coefficient (r) between reflectance/FDR and Chl-*a*/TSS with regard to each waveband. It is clear that FDR obviously improved correlation with Chl-*a* and TSS; moreover, spectra reflectance and absorption features were also enhanced (Fig. 3.1b). An NIR/red algorithm developed by Han et al. (1997) was introduced for comparison of the RSI selected wavebands and accuracy. The NIR/red model showed a higher R^2 and lower RMSE than the single waveband models. However, based on the regression between the reflectance of each waveband and Chl-*a* and TSS, the RSI model selected the R719/R662 ratio as the best band combination, which enhanced the performance of ratio model, giving the highest R^2 value of 0.72 for Chl-*a*. The R717/R630 ratio was the best band combination for TSS, with an R^2 of 0.52 (Fig. 3.2a, b). A three-band semi-analytical algorithm for estimating Chl-*a* concentration was conducted, as a previous study suggested (Gitelson et al., 2007), and the optimal wavebands of model were tuned according to the optical properties of the water bodies. Bands 660, 703, and 740 nm were final selected for the three-band model with an R^2 of 0.71 and RMSE of 29.32. For another algorithm introduced in a previous study, the NDCI was evaluated using remote sensing reflectance

R_{rs} at an absorption peak of 665 nm (i.e., $R_{rs}(665)$), which is closely related to absorption by Chl-*a* pigments and a reflectance peak of 708 nm (i.e., $R_{rs}(708)$), which was sensitive to variations in Chl-*a* concentration in water, with a result of an R^2 of 0.60 and an RMSE of 28.82. For the NDSI model, bands 719 and 663 nm were the best combination for estimating Chl-*a* ($R^2 = 0.64$), and bands 704 and 698 nm were the best combination for TSS ($R^2 = 0.55$) (Fig. 3.2c, d). The results showed the lowest RMSECV in the RSI model for Chl-*a* (24.14) and in the NDSI model for TSS (8.48). Among the models, the RSI or the NDSI showed higher R^2 values and lower RMSECV values than those of the two types of single-band models in the estimation of both Chl-*a* and TSS.

Table 3.2. Regression models used to estimate Chl-*a* and TSS concentrations with two spectral data types (reflectance and FDR) and two spectral indices (RSI and NDSI).

Parameter	**Spectral index**	**Model**	R^2	**RMSE**
Chl-*a*	Reflectancce	Chl-$a = 0.0004 \times R_{730} + 0.0396$	0.14	51
	FDR	Chl-$a = 1 \times 10^{-5} \times R_{705} - 0.0004$	0.54	51.01
	NIR/red (Han et al. (1997))	Chl-$a = 94.748 \times R_{705}/R_{670} - 88.897$	0.6	28.78
	Three-band (Gitelson et al. (2003))	Chl-$a = 0.0036 \times (R^{-1}{}_{660} - R^{-1}{}_{703}) \times R_{740} - 0.0665$	0.71	29.32
	NDCI (Mishra et al. (2012))	Chl-$a = 253.16 \times (R_{rs708} - R_{rs665})/(R_{rs708}+R_{rs665}) + 36.535$	0.6	28.82
	RSI	Chl-$a = 119.27 \times R_{719}/R_{662} - 88.052$	0.72	24.14
	NDSI	Chl-$a = 253.16 \times (R_{719} - R_{663})/(R_{719} + R_{663}) + 36.535$	0.64	27.19
TSS	Reflectancce	$\text{TSS} = 0.0009 \times R_{722} + 0.0501$	0.05	14.81
	FDR	$\text{TSS} = 5 \times 10^{-5} \times R_{704} - 0.0003$	0.46	14.83
	RSI	$\text{TSS} = 31.419 \times R_{717}/R_{630} - 17.913$	0.52	8.73
	NDSI	$\text{TSS} = 300.45 \times (R_{704} - R_{698})/(R_{704} + R_{698}) + 6.3868$	0.55	8.48

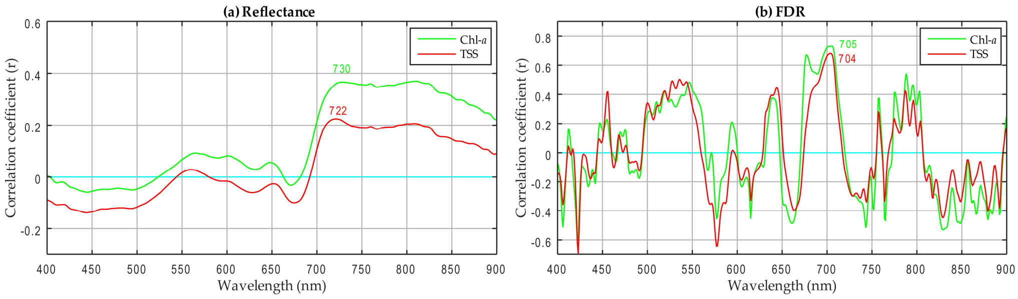

Figure 3.1. Correlation coefficients (*r*) between water quality parameters (Chl-*a* and TSS) at each wavelength: (a) reflectance; (b) FDR.

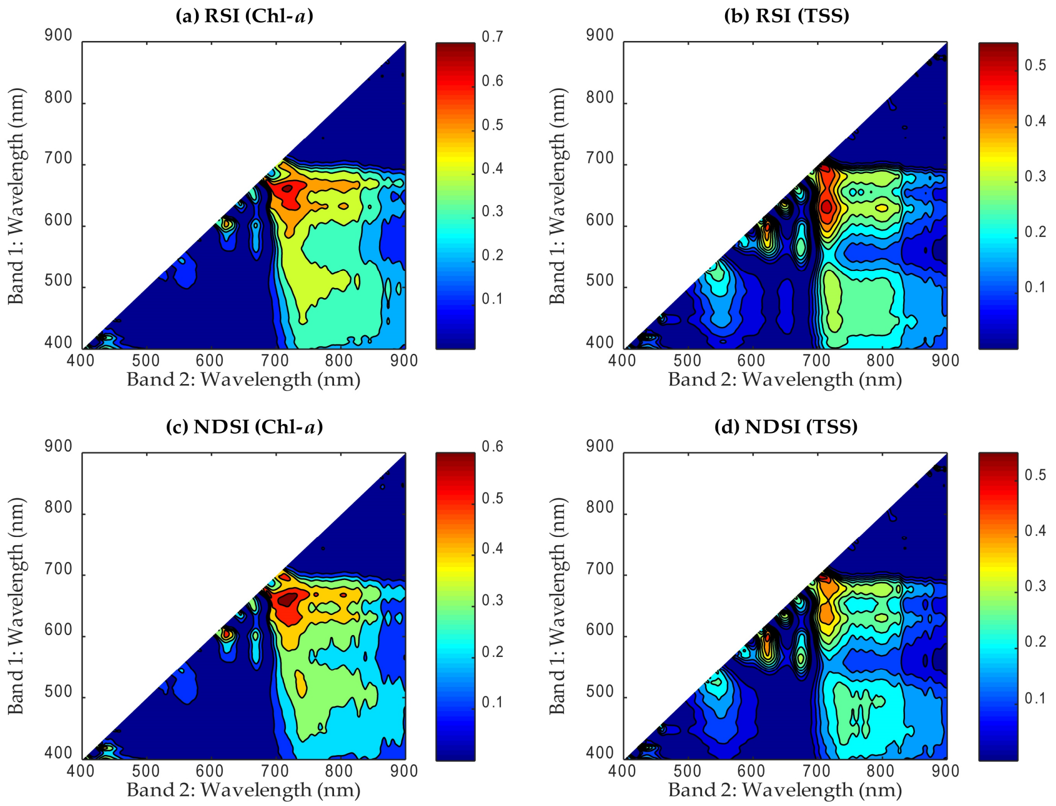

Figure 3.2. Distributions of R^2 between two wavebands using RSI (a) Chl-*a*; (b) TSS and NDSI (c) Chl-*a*; (d) TSS.

3.4.3. FS–PLS and ISE–PLS Models

Calibration and cross validation results between reflectance/FDR spectra and Chl-*a*/TSS using FS–PLS and ISE–PLS are shown in Table 3.3. The results showed that the optimum NLV ranged between 4 and 8 in FS–PLS and between 5 and 11 in ISE–PLS, which was determined by the LOO cross validation based on the lowest RMSECV. The RPD ranged between 1.22 and 1.32 (low accuracy) in FS–PLS and between 1.45 and 7.44 (excellent accuracy) in ISE–PLS. In particular, the selected number of wavebands and the percentage to full spectrum (that is, selected wavebands number/all (n = 501) × 100%) were calculated to evaluate the informative wavebands for ISE–PLS. Results showed the selected wavebands number ranged between 9 and 85, and the percent ratio ranged between 1.80 and 16.97. Overall, for Chl-*a*, ISE–PLS using FDR showed the highest R^2, highest RPD, and lowest RMSECV (R^2 = 0.98, RMSECV = 6.15, RPD = 7.44); NLV = 11, and 85 wavebands were selected. Similarly, with respect to TSS, ISE–PLS using FDR showed the highest R^2, highest RPD and lowest RMSECV (R^2 = 0.97, RMSECV = 1.91, RPD = 6.64); NLV = 11, and 42 wavebands were selected.

Table 3.3. Optimum NLV, R^2 and RMSECV using the LOO method in FS–PLS and in ISE–PLS using the entire dataset (n = 36), with the residual predictive deviation, the number of selected wavebands and the percent ratio with respect to the full spectrum (i = 501).

Parameter	Spectral Data Type	Regression	Calibration			Cross Validation			Selected Wavebands Number	Selected Wavebands (%)
			NLV	R^2	RMSEC	R^2	RMSECV	RPD		
Chl-*a*	Reflectance	FSPLS	4	0.59	29.26	0.41	35.44	1.28		
	Reflectance	ISEPLS	6	0.7	25.01	0.6	29.27	1.55	9	1.8
	FDR	FSPLS	8	0.99	3.25	0.43	35.15	1.32		
	FDR	ISEPLS	11	1	1.14	0.98	6.15	7.44	85	16.97
TSS	Reflectance	FSPLS	6	0.61	7.87	0.35	10.36	1.22		
	Reflectance	ISEPLS	5	0.62	7.76	0.53	8.73	1.45	13	2.59
	FDR	FSPLS	5	0.93	3.39	0.4	9.98	1.27		
	FDR	ISEPLS	11	1	0.84	0.97	1.91	6.64	42	8.38

FDR = first derivative reflectance; NLV = number of latent variables; RMSEC = root mean square error from calibration; RMSECV = root mean square error from cross validation; RPD = the residual predictive deviation.

The relations between observed and predicted Chl-*a* and TSS are shown in Fig. 3.3. The data in this figure were used to evaluate goodness of fit in the FS–PLS and ISE–PLS models. Comparisons between the FS–PLS and ISE–PLS models were presented in combination with the R^2 and RMSE from the cross validation listed in Table 3.3. For Chl-*a*, the ISE–PLS using FDR showed a higher R^2 and lower RMSECV. The scatter distribution also showed a better linear relation, which can be judged by the red dots clustered along the 1:1 line in Fig. 3.3b. Similarly, with respect to TSS, the ISE–PLS model using FDR showed better results than the others (Fig. 3.3d, red dot). However, both red and green dots clustered vertically, particularly in Fig. 3.3a, c, showing a large variation in the predicted values and nearly no variation in the observed values, indicating that plenty of observed Chl-*a* and TSS samples had low concentrations. This vertical clustering also indicates the FS–PLS and ISE–PLS using reflectance had lower predictive abilities than using FDR.

The selected wavebands in ISE–PLS using the reflectance and FDR datasets are shown in Fig. 3.4. In the reflectance dataset, the selected wavebands were primarily in the red wavelengths (650–680 nm) for Chl-*a*. For TSS, the selected wavebands were in green wavelengths (560 nm), red wavelengths (620–630 nm) and red-edge wavelengths (720 nm). In the FDR datasets, a cluster of wavebands focus on the red region (670–680 nm, 690–710 nm) for Chl-*a*, and wavebands were also selected from other regions: blue (around 410), green (around 490 nm, 510 nm), red (around 603, 615), and the NIR region between 820 nm and 900 nm. Similarly, more wavebands were selected for TSS using FDR than using reflectance, especially in the red (around 620 nm, 680 nm, and 700 nm) and NIR (around 730 nm) regions.

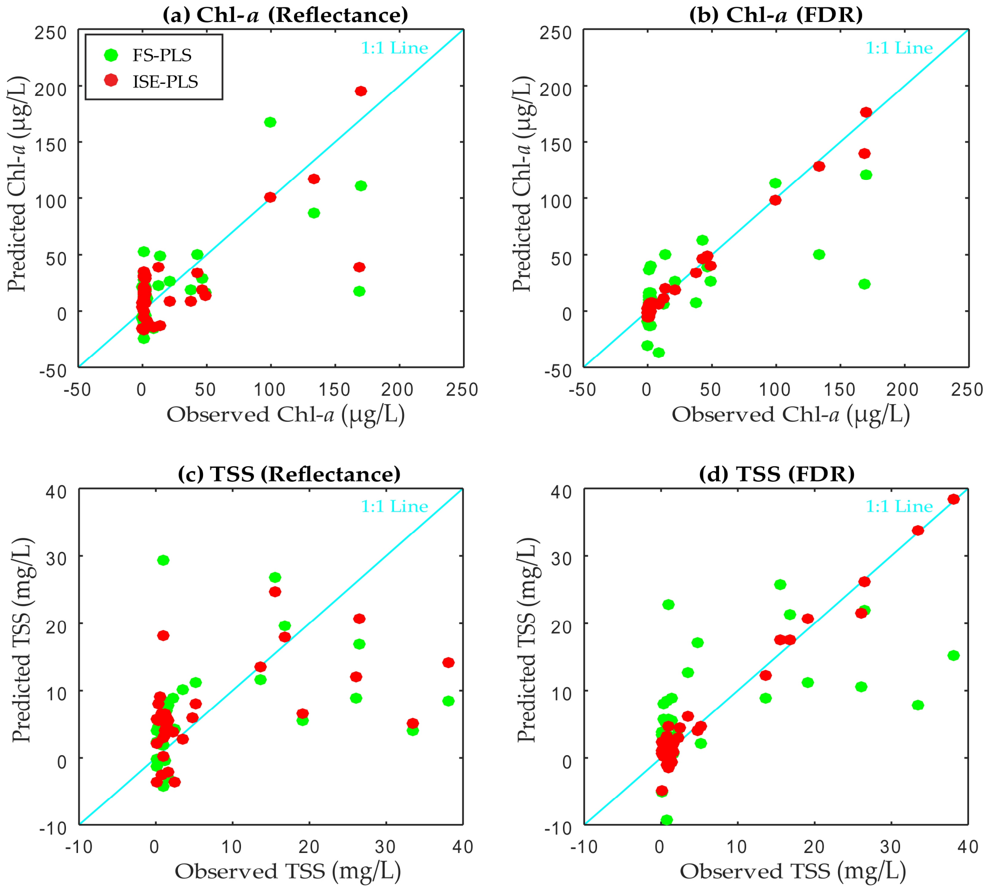

Figure 3.3. Relations between measured and cross-validated prediction values of Chl-*a* (**a**) Reflectance; (**b**) FDR and TSS; (**c**) Reflectance; (**d**) FDR using FS–PLS and ISE–PLS.

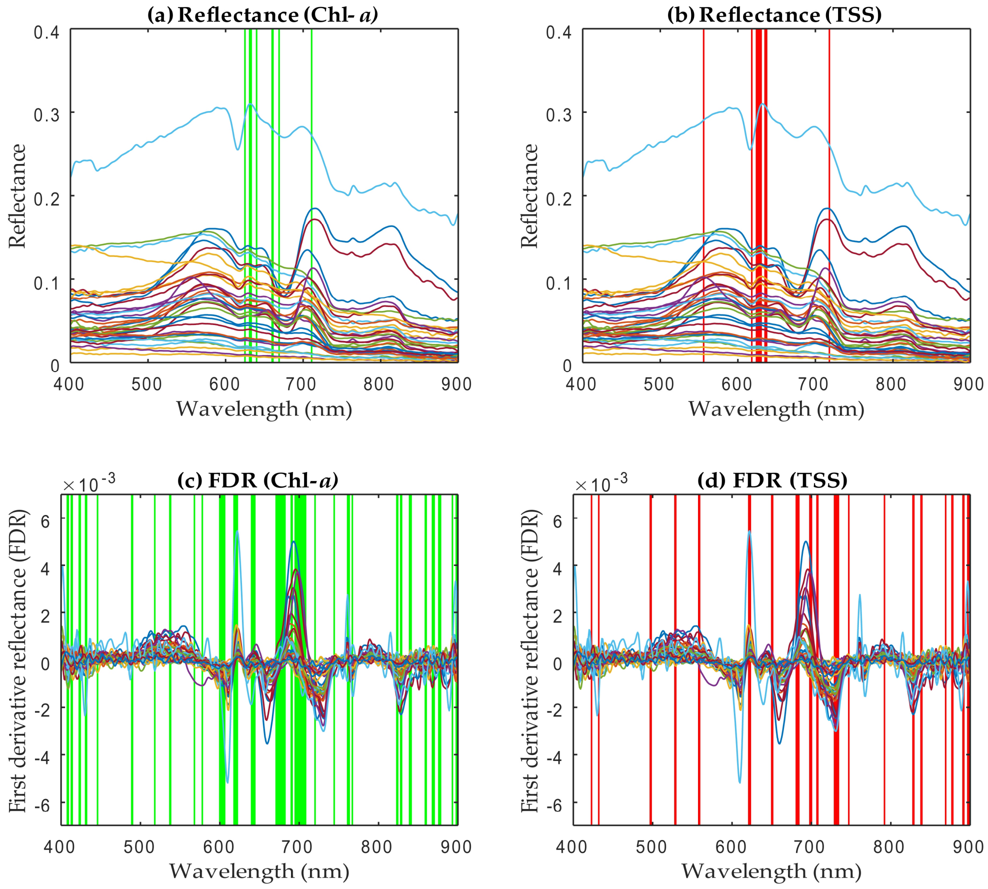

Figure 3.4. Selected wavebands in ISE–PLS using reflectance or FDR datasets (n = 36) to estimate: (**a**) and (**c**) Chl-*a*; (**b**) and (**d**) TSS. Green bars = Chl-*a*; red bars = TSS.

3.5. Discussion

3.5.1. Evaluation of the Predictive Abilities of Simple Linear Regression Models

In the present study, models established by single waveband and two waveband combinations were compared using PLS. For single waveband models, FDR showed a better R^2 than smoothed reflectance both for Chl-*a* and TSS, indicating that the accuracy can be improved by enhancing the features of absorption and reflectance from the smoothed reflectance. However, all single waveband models showed poor accuracy for estimating both Chl-*a* and TSS concentrations. According to previous research, single band focus on 670–750 nm is better at determining TSS concentrations (Nechad et al., 2010), especially in turbid water. Single band focus showed no predictive ability in that research, and simple linear regression using two wavebands combinations showed poor accuracy for TSS, which may indicate that TSS is difficult to detect using single-band or two-band combinations in relatively clear water; as shown in our results, most observed TSS values were low. The three-band model was successfully used for Chl-*a* retrieval in turbid water bodies (Gitelson, et al., 2008; Gitelson et al., 2007). As for this research, the optimal spectral bands selected from the iterative band tuning are in accord with the previous research (Dall'Olmo et al., 2003); however, even the result shows a considerable R^2, but the relatively high RMSE may indicate a low accuracy model, which may be attributed to different compositions of optically active constituents (Chl-*a*, tripton, CDOM) (Gitelson et al., 2007). The NDCI is a special case of the NDSI: two bands of NDCI are determined by the reflectance peak and spectral absorption peak, and the normalizing of two bands reflectance can eliminate uncertainties in the estimation of R_{rs} (Mishra & Mishra, 2012). As a comparison, the result of the NDSI has a slight improvement with an R^2 of 0.64 and an RMSE of 27.19 than the NDCI with an R^2 of 0.60 and an RMSE of 28.82, which may indicate that a combination of wavebands at 719 and 663 nm in the NDSI can better reflect the Chl-*a* variations in this research area. Among all tested combinations of the RSI and the NDSI, the best R^2 values were obtained using the NIR waveband (719 nm) and the red region (662 nm for the RSI, 663 nm for the NDSI) to estimate Chl-*a* concentrations, which agrees with the findings of other research. In most available research on the measurement of chlorophyll content in water, the absorption trough is located at near 670 nm, caused by absorption of Chl-*a* (Han &

Rundquist, 1997; Huang et al., 2010) and the reflectance peak near 710 nm, caused by the fluorescence of Chl-*a* (Gitelson A. , 1992; Bennet & Bogorad, 1973; Ma et al., 2007) On account of these characteristics, the two waveband models, particularly the NIR/red ratio, have been widely used for Chl-*a* retrieval, and a variety of algorithms have been based mainly on the ratio of reflectance peak (about 710 nm) to reflectance trough (about 670 nm) (Han & Rundquist, 1997; Mittenzwey et al., 1991). Similarly, in the present study two wavebands, from the NIR and red regions respectively, were selected by the NDSI, confirming the water body reflection characteristics.

3.5.2. Evaluation of the Predictive Abilities of FS–PLS and ISE–PLS

As we expected, the PLS models exhibited better predictive abilities than models that use single wavebands or the index-based (RSI and NDSI) approaches, which shows the PLS method is potentially useful in retrieval of inland water quality parameters (Song et al., 2013; Ryan & Ali, 2016). In our PLS analyses, results using ISE–PLS models with the FDR dataset showed higher R^2 and lower RMSECV values than those of the reflectance dataset. These results are consistent with the research of Han and Rundquitst (1997), who noted that FDR was better correlated with chlorophyll concentration than raw reflectance, and that random noise and the effects of suspended matter could be reduced by FDR (Song et al., 2013). After eliminating outliers and useless predictors, ISE–PLS calibrated more potential models than FS–PLS, both for Chl-*a* and TSS, with the wavelengths relevant to water quality. As a consequence, predictive ability was further enhanced, which is reflected in the results of evaluation indices. PLS-based waveband selection greatly improved predictions for both Chl-*a* (R^2 from 0.43 to 0.98, RMSECV from 35.15 to 6.15, RPD from 1.32 to 7.44) and TSS (R^2 from 0.40 to 0.97, RMSECV from 9.98 to 1.91, RPD from 1.27 to 6.64). The PLS models in combination with wavelength selection had an improved performance also supported by other previous research (Kawamura et al., 2008; Forina et al., 2004; Chen et al., 2007). However, the R^2 for Chl-*a* using ISE–PLS reached 0.98, a result that does not rule out the possibility of overfitting; therefore, the solution method for this condition should be the subject of additional research and validation.

3.5.3. Importance of Selected Wavebands in ISE–PLS

Our results showed 16.97% of all available wavelengths that were selected for predicting Chl-*a* and 8.38% were also selected for predicting TSS by ISE–PLS, which indicates that less than 20% of the waveband information from field hyperspectral data contributes to the prediction for water quality parameters (Chl-*a* and TSS) and over 80% were redundant. In the reflectance dataset, wavebands primarily in the red wavelengths were selected: between 630 and 710 nm for Chl-*a*; for TSS, 560 nm, 620–630 nm, and 720 nm. In the FDR dataset, the selected wavebands for estimating both Chl-*a* and TSS involved more regions than the reflectance dataset. Nevertheless, similar wavelengths in the visible and NIR regions were selected; blue (410 nm), green (approximately 490 nm, 510 nm), and red (approximately 603 nm, 615 nm) for Chl-*a*; and blue (approximately 420 nm), green (approximately 500 nm), red (approximately 620 nm, 680 nm and 700 nm), and NIR (approximately 730 nm) for TSS. Intensive absorption by Chl-*a* resulted in reflectance troughs around 440 and 670nm (Fig. 3.4a) (Yacobi et al., 2011). Low absorption of algal pigments or the scattering of phytoplankton cells and inorganic suspended materials might cause the reflectance peak near 570 nm (Huang et al., 2010). The reflectance spectrum peak near 700 nm had a strong correlation with Chl-*a* concentration (Gitelson A. , 1992; Gitelson et al., 1993). Several previous studies of inland water quality also proved these wavelengths have the potential to predict Chl-*a* and TSS concentrations (Hu et al., 2011; Gons, 1999). This study brings obvious evidence that the ISE-PLS model may be considered as a unified approach for remote quantification of constituent concentrations in water quality assessment. Using this method, more informative wavebands can be selected from hundreds of hyperspectral wavebands, which indicates the accuracy and efficiency can be enhanced by ISE-PLS when it comes to using hyperspectral sensors in satellites with a high temporal and spatial resolution to monitor relatively small area inland water quality in the future.

3.6. Conclusions

The present study develops models for estimating Chl-*a* and TSS concentrations in irrigation ponds using water surface reflectance spectral data. Our results show that PLS regression analysis has high potential for predicting Chl-*a* and TSS based on field hyperspectral measurements, and that ISE wavebands selection in combination with PLS regression analysis can enhance predictive ability. Chl-*a* and TSS concentrations were estimated with high accuracy by using ISE-PLS, which explains 98% of the variance for Chl-*a* and 97% of the variance for TSS. The important wavebands for estimating Chl-*a* and TSS using ISE–PLS represented 16.97% and 8.38%, respectively, of all 501 wavebands over the 400–900 nm range. The selected wavebands approximately match the absorption peaks published by previous researchers. Compared to the estimation of water quality parameters by satellite sensors such as MODIS, ISE–PLS selected more informative wavebands, especially the wavelength at approximately 700 nm. These results provide useful insights for future analyses on the assessment of water quality in irrigation ponds, especially when using satellite imagery.

Chapter 4: Validation of new chlorophyll-*a* algorithm in the Seto Inland Sea using ISE-PLS method

4.1. Introduction

The Seto Inland Sea is an approximately 23,000 km^2 semi-enclosed coastal sea in western Japan, with an average depth of 38 m. This sea is well-known as one of the world's most productive fisheries due to its abundance of fish and variety of fish species (Hashimoto et al., 1997). Approximately 35 million people live around the Seto Inland Sea, bringing increased industrialization and urbanization that has made the Seto Inland Sea one of Japan's most industrialized regions (Yoshie et al., 2011). However, productivity of fisheries is sensitive and thus vulnerable to anthropogenic stress. Eutrophication of coastal waters has affected fishing and other activities by contributing to harmful algal blooms (HABs), also known as red tides. HABs frequently occurred in the Seto Inland Sea during a period of high economic growth in the 1970s (Nishijima et al., 2016). Although HABs have decreased from about 300 cases per year in 1976 to about 100 cases per year more recently (Yamamoto, 2003), severe damage to fisheries and significant economic losses due to HABs are still occurring (Imai et al., 2006). Therefore, monitoring HABs is vital for managing the fisheries industry and ensuring sea water quality.

The scientific community and various agencies monitor HABs to manage and control them. Many studies have conducted HABs observation, but most have relied on conventional *in situ* ship surveys and buoy stations (Yunus et al., 2015). For each observation, water samples must be collected and analyzed under controlled lab environments, which is costly and time consuming. Additionally, the spatial scale of such surveys is limited. Remote sensing has been widely applied to monitor, in real time, various ocean environment factors on a large scale using spaceborne or airborne instruments. Having a high degree of spatial and temporal coverage over a large scale is convenient for monitoring HABs. Several studies have shown distributions of HABs using satellite imagery and Chl-*a* concentration measurements (Noh et al., 2018; Zhao and Ghedira, 2014). Chl-*a* concentration in water is a major indicator of a trophic state and oceanic Chl-*a* concentration is the most common property characterizing first trophic levels in marine environments (Wang and Liu, 2005). Chl-*a* acts as a link between nutrient concentration and algal production (Gholizadeh et al., 2016), therefore, it can be used as a proxy to evaluate HABs.

Earlier studies indicated bio-optical model was usually used to estimate Chl-*a* concentration in water (Katlane et al., 2012). Several satellite sensor algorithms developed to estimate ocean and coastal Chl-*a* using the remote sensing reflectance (R_{rs}), such as the ocean chlorophyll 2-band (OC2), ocean chlorophyll 3-band (OC3), and ocean chlorophyll 4-band (OC4) algorithms used for the standard Sea-viewing Wide Field-of-view Sensor (SeaWiFS) Chl-*a* product (O'Reilly et al., 1998; O'Reilly et al., 2000). The strong Chl-*a* absorption in red bands and diminishing Chl-*a* absorption and increasing water absorption in near infrared (NIR) bands (Moses et al., 2011) yields a band ratio between the NIR and red bands that has frequently been used to estimate Chl-*a* concentrations (Gitelson et al., 1993; Mittenzwey et al., 1992; Sakuno et al., 2014), Han and Rundquist (1997) found that the ratio of reflectance at 705 nm (NIR) to reflectance at 670 nm (red) correlated well with Chl-*a* concentration in a turbid reservoir.

Additionally, a three-band semi-analytical reflectance model can be used to assess Chl-*a* concentration by taking advantage of the red and NIR regions' absorption characteristics. This model was originally developed to estimate pigment content in higher plant leaves (Dall'Olmo et al., 2003).

Partial least squares (PLS) regression, a statistical method developed by Wold (1966), is an efficient tool for multivariate modeling that is increasingly used to handle high-dimensional hyperspectral data (Kawamura et al., 2008; Song et al., 2013). Its potential application for water quality quantification has been tested (Song et al., 2012), and the iterative stepwise elimination PLS (ISE-PLS; Boggia et al., 1997), which combines PLS and a wavelength selection function, has proven effective at estimating Chl-*a* in inland case II water (Wang et al., 2017). However, ISE-PLS has not been tested in coastal waters or compared with traditional algorithms.

Our objectives are: (1) to develop models to estimate Chl-*a* using *in situ* hyperspectral data; (2) to evaluate traditional empirical and semi-analytical algorithms in Seto Inland Sea; and (3) to evaluate the ISE-PLS method's accuracy in coastal waters.

4.2. Materials and methods

4.2.1. Study area

The study area is in the central part of the Seto Inland Sea near the city of Fukuyama as shown in Fig. 1.2. We selected six sites as sampling stations, which are described in Table 1.2.

4.2.2. Data collection and pre-processing

We conducted 12 ship-based surveys were conducted between December 16, 2015, and September 7, 2017, and obtained 59 datasets from six stations. We performed *in situ* measurements of water-leaving reflectance (R_L) using a MS-720 (Eiko Co. Ltd, Tokyo, Japan) spectrometer, with a spectral range of 350–1050 nm and a spectral interval of 3.3 nm. We recalculated the spectral interval to 1 nm when exporting data. We gathered spectral readings approximately 1 m above the water surface, with a probe field angle of 25°, between 9:00 and 11:00 under clear sky conditions. We measured Chl-*a* using a Hydrolab DS5 (Hach, Loveland, USA) multiparameter data sonde with sensors for measuring Chl-*a* and other water quality parameters. In this study we used Chl-*a* data from just beneath the water surface.

With respect to spectral data, we identified reflectance ranges of 325–399 nm and 901–1075 nm as noise and removed them. We then smoothed the spectral data using a Savitzky-Golay filter with 15 smoothing points. We also computed the first derivative reflectance (FDR) and compared it with the original water-leaving reflectance for Chl-*a* estimation

4.2.3. Algorithms

In this study we used the newest OC algorithms (version 6; Hu et al., 2012), which was formulated as a fourth-order polynomial with five coefficients. The modified OC algorithms yielded better statistical agreement between model data and Chl-*a* than the first version OC algorithm (O'Reilly et al., 1998), which was a modified cubic polynomial relationship between Chl-*a* and a ratio of R_{rs}.

R_{rs} (λ) can be represented by the relationship with water-leaving reflectance R_L (λ) as follows (Oyama et al., 2009):

$$\mathrm{R}_{rs}(\lambda) = \pi \mathrm{R}_L(\lambda) \tag{4.1}$$

where λ is wavelength. The version 6 OC algorithms use a fourth-order polynomial equation that can be written as:

$$log_{10}(Chl\_a) = a_0 + a_1R + a_2R^2 + a_3R^3 + a_4R^4 \tag{4.2}$$

where R and the coefficients a_0, a_1, a_2, a_3, and a_4 are different in the OC2, OC3, and OC4 algorithms. OC2 uses the blue/green ratio $R_{rs}(490)/R_{rs}(555)$ and R is described as R = $\log_{10}(\frac{R_{rs}\,(490)}{R_{rs}\,(555)})$. The coefficients are given as a_0 = 0.2511, a_1 = -2.0853, a_2 = 1.5035, a_3 = -3.1747, and a_4 = 0.3383. OC3 uses a three-band formulation with a maximum of R_{rs} band ratios $R_{rs}(443)/R_{rs}(555)$ and $R_{rs}(490)/R_{rs}(555)$, and R is expressed as R = $\log_{10}(\frac{R_{rs}\,(443)>R_{rs}\,(490)}{R_{rs}\,(555)})$ with coefficients of a_0 = 0.2515, a_1 = -2.3798, a_2 = 1.5823, a_3 = -0.6372, and a_4 = -0.5692. Similarly, OC4 uses the maximum of three R_{rs} ratios—R_{rs} (443)/R_{rs} (555), R_{rs} (490)/R_{rs} (555) and R_{rs} (510)/R_{rs} (555)—to build the formulation, with R expressed as R = $\log_{10}(\frac{R_{rs}\,(443)>R_{rs}\,(490)>R_{rs}\,(510)}{R_{rs}\,(555)})$ and coefficients a_0 = 0.3272, a_1 = -2.9940, a_2 = 2.7218, a_3 = -1.2259, and a_4 = -0.5683. The OC4 algorithm has been considered a standard method for satellite detection of HABs over global waters (Stumpf, 2001; Tomlinson et al., 2004).

The three-band model uses the NIR and red wavebands and is formulated as (Dall'Olmo et al., 2003; Dall'Olmo and Gitelson, 2005):

$$\mathrm{Chl\_}a \propto (R_{\lambda 1}^{-1} - R_{\lambda 2}^{-1}) \times R_{\lambda 3} \tag{4.3}$$

where $R_{\lambda i}$ is the reflectance at a wavelength of λ_i nm. We found the optimal spectral ranges for these wavelengths to be, λ_1 = 660–670 nm, which is maximally sensitive to absorption by Chl-*a*; λ_2 = 690–720 nm, which is minimally sensitive to absorption by Chl-*a*; and λ_3 = 720–750 nm, which is minimally affected by absorption by any constituent (Dall'Olmo and Gitelson, 2005; Gitelson et al., 2008). We expected to find the optimal spectral ranges of λ_1, λ_2, and λ_3 for Chl-*a* estimation by spectrally tuning the conceptual model using a stepwise technique. First, we set λ_2 and λ_3 to 700 nm and 750 nm, respectively, and then linearly regressed using all available bands and Chl-*a* to obtain the first estimate of λ_1, with which there was a high correlation coefficient (*r*). After we fixed λ_1, we set λ_2 as an unknown waveband and linearly regressed to find an optimal λ_2 based on the best r value using the reflectance corresponding to a fixed λ_1 and an assumed λ_3. Analogously, we confirmed the optimal λ_3 using the reflectance corresponding with fixed λ_1 and λ_2 values.

The NIR/red two-band model has been widely used to retrieve Chl-*a* concentrations in turbid productive waters to identify phytoplankton blooms (Gitelson, 1992). This model is formulated as follows:

$$\mathrm{Chl\_}a \propto R_{\lambda 1}^{-1} \times R_{\lambda 2} \tag{4.4}$$

where λ_1 is in the red region and λ_2 is in NIR region. We tuned the model to select the optimal NIR and red bands for Chl-*a* retrieval in this research area and compared its accuracy with the previous model using wavelengths of 705 nm in the NIR region and 670 nm in the red region.

PLS is useful for handling many descriptors even when co-linearity and noise in the model building regression are present (D'Archivio et al., 2014). The standard PLS regression equation can be expressed as follows:

$$y = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_i x_i + \varepsilon \tag{4.5}$$

where *y* is the response variable that represents Chl-*a*, *xi* is the predictor variable representing spectral data such as R_L or FDR values for spectral bands 1 to *i* (400–900 nm), β_i is the estimated weighted regression coefficient, and *ε* is the error vector. In the PLS model, the original predictor variables (X) are projected onto a small number of orthogonal latent variables to simplify their relationships with response variables (Y) (Y; Li and He, 2003). We selected the optimal number of latent variables (NLV) in the final model using the leave-one-out (LOO) cross-validation method with a minimum value of the root mean squared error (RMSE), which is calculated as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(y_i - y_p\right)^2}{n}} \tag{4.6}$$

where y_i and y_p represent sample i's measured and predicted Chl-*a*, respectively, and *n* is the number of samples in the dataset (*n* = 59).

The ISE-PLS uses a model-wise elimination technique (Boggia et al., 1997) that permits the removal of useless descriptors to improve predictive performance. This process is based on the importance of the predictor z_i, which is defined as:

$$z_i = \frac{|\beta_i| s_i}{\sum_{i=1}^{I} |\beta_i| s_i} \tag{4.7}$$

where s_i is the standard deviation of predictor x*i*. PLS modeling uses all available wavebands (501 bands between 400 and 900 nm). Predictors are then evaluated based on the value of the importance of predictor z_i. The predictor with minimum importance is eliminated in each elimination cycle and the remaining predictors are used to recalibrate the model (Forina et al., 2003). Finally, a model with maximum predictive ability is selected using the minimum RMSE value from the cross-validation.

4.2.4. Evaluation of predictive ability

We used the coefficient of determination (R^2) and RMSE to evaluate the predictive ability of empirical and semi-analytical algorithms such as OC, three-band, and two-band model algorithms. Higher R^2 values and a lower RMSE indicate better Chl-*a* estimation performance. To evaluate the ISE-PLS predictive ability, we used R^2 and RMSE from the LOO cross-validation in the final model. Additionally, we introduced the residual predictive deviation (RPD), which is defined as the ratio of the standard error of the prediction to the standard deviation, as the evaluating indicator. RPD can be expressed as RPD = SD/RMSE (Williams, 2001). As shown in a previous study by Chang and Laird (2002), an RPD > 2 indicates a model with good predictive ability, 1.4 < RPD < 2 indicates moderately good model in need of some improvement, and an RPD < 1.4 means the model has no predictive ability.

We performed all data handling and regression analyses using Matlab software ver. 8.6 (MathWorks, Sherborn, MA, USA).

4.3. Results

4.3.1. Chl-*a* characteristics

Table 4.1 shows Chl-*a* concentration descriptive statistics from this study, including stations, number of samples, minimum (Min), maximum (Max), mean, standard deviation (SD) and coefficient of variation (CV).

Table 4.1. Chl-*a* concentration descriptive statistics.

Stations	N	Min	Max	Mean	SD	CV
1	12	0.83	4.2	2.73	0.95	0.35
2	12	1.06	6.72	3.82	2.15	0.56
3	12	1.72	7.84	4.5	1.71	0.38
4	12	2.31	14.33	8.13	4.54	0.56
5	6	1.75	5.46	3.92	1.25	0.32
6	5	1.2	8.74	4.41	2.84	0.64
Total	59	0.83	14.33	4.67	3.11	0.67

N, number of samples; SD, standard deviation; CV, coefficient of variation.

4.3.2. Comparison of empirical and semi-analytical models

We used several empirical and semi-analytical models for Chl-*a* retrieval, the results of which are shown in Table 4.2. We initially used three standard empirical algorithms, OC2, OC3, and OC4. The first row of Fig. 4.1 shows scatter plots between *in situ* measured Chl-*a* and Chl-*a* derived from OC models. The results show a linear relationship between measured and modelled Chl-*a* for all three OC algorithms, with poor R^2 values (0.36, 0.31, and 0.30, respectively for OC2, OC3, and OC4). In addition, results of all three OC algorithms underestimate Chl-*a*. In view of the sensor differences between the hyperspectral spectrometer used for *in situ* measurements and the SeaWiFS satellite sensor, we recalculated the parameters for equation (4.2) by model recalibration using *in situ* Chl-*a* and R_{rs}, which was in accordance with specified OC algorithm wavebands. The second row of Fig. 4.1 shows scatter plots between recalibrated OC algorithms and Chl-*a*. The R^2 values for all three OC algorithms were slightly improved (0.39, 0.36, and 0.35 respectively for OC2, OC3, and OC4), and scattered points were close to the 1:1 line. For both the standard and recalibrated OC models, the OC2 algorithm performed better than

OC3 and OC4 for Chl-*a* retrieval in this study; however, its predictive ability remains poor due to its low R^2 value.

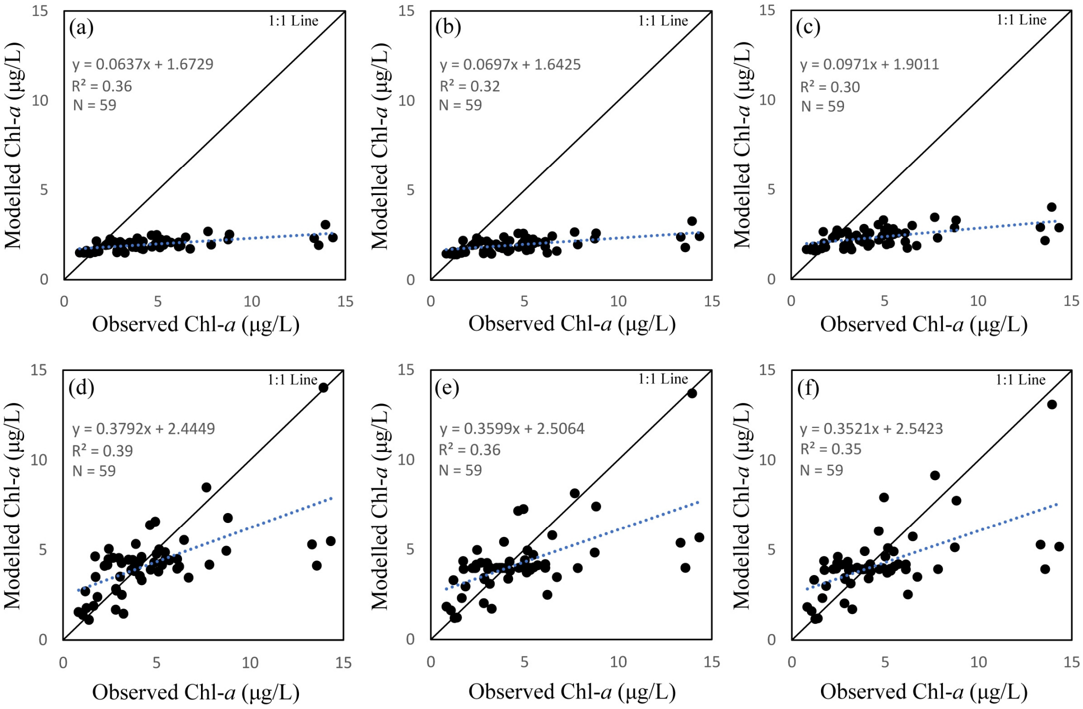

Figure 4.1. Correlation between observed and modelled chlorophyll *a* (Chl-*a*) using ocean chlorophyll (OC) algorithms. The first row shows results using standard coefficients (a) OC2, (b) OC3, and (c) OC4. The second row shows results from recalibrated OC models using the dataset of this study (d) OC2, (e) OC3, and (f) (OC4).

Table 4.2. Regression models used to estimate Chl-*a* concentrations.

Algorithms	Results equation	Bands combination (*R*), Coefficient *a*, and Intercept *b*	R^2	RMSE
OC2	$Chl\_a = 10^{a_0+a_1R+a_2R^2+a_3R^3+a_4R^4}$	$R = \log_{10}(\frac{R_{rs}(490)}{R_{rs}(555)})$ *a* = [0.2511 -2.0853 1.5035 -3.1747 0.3383]	0.36	3.96
OC3	$Chl\_a = 10^{a_0+a_1R+a_2R^2+a_3R^3+a_4R^4}$	$R = \log_{10}(\frac{R_{rs}(443)>R_{rs}(490)}{R_{rs}(555)})$ *a* = [0.2515 -2.3798 1.5823 -0.6372 -0.5692]	0.32	3.95
OC4	$Chl\_a = 10^{a_0+a_1R+a_2R^2+a_3R^3+a_4R^4}$	$R= \log_{10}(\frac{R_{rs}(443)>R_{rs}(490)>R_{rs}(510)}{R_{rs}(555)})$ *a* = [0.3272 -2.9940 2.7218 -1.2259 -0.5683]	0.30	3.66
Recalibrated OC2	$Chl\_a = 10^{a_0+a_1R+a_2R^2+a_3R^3+a_4R^4}$	$R = \log_{10}(\frac{R_{rs}(490)}{R_{rs}(555)})$ *a* = [-8942.6 -2053.3 -100.25 -3.8257 0.5738]	0.39	2.65
Recalibrated OC3	$Chl\_a = 10^{a_0+a_1R+a_2R^2+a_3R^3+a_4R^4}$	$R = \log_{10}(\frac{R_{rs}(443)>R_{rs}(490)}{R_{rs}(555)})$ *a* = [5204.7 -461.22 -41.033 -4.4207 0.5491]	0.36	2.50
Recalibrated OC4	$Chl\_a = 10^{a_0+a_1R+a_2R^2+a_3R^3+a_4R^4}$	$R = \log_{10}(\frac{R_{rs}(443)>R_{rs}(490)>R_{rs}(510)}{R_{rs}(555)})$ *a* = [-30610 -4098 -57.405 -0.1942 0.5933]	0.35	2.53

Three-band	$Chl\_a = aR + b$	$R = (R(664)^{-1} - R(695)^{-1}) \times R(736)$ $a = 85.096$ $b = 7.371$	0.46	2.28
NIR/red	$Chl\_a = aR + b$	$R = R(705) \times R(670)^{-1}$ $a = 0.0044$ $b = 0.8863$	0.17	4.88
NIR/red tuning	$Chl\_a = aR + b$	$R = R(693) \times R(666)^{-1}$ $a = 66.633$ $b = -59.755$	0.39	2.40

OC2, ocean chlorophyll-2; OC3, ocean chlorophyll-3; OC4, ocean chlorophyll-4; NIR, near-infrared.

The three-band and two-band algorithms were both based on the NIR region which has high absorption by water and the red region which has high absorption by Chl-*a*. Fig. 4.2 shows the three-band algorithm tuning process. The optimal λ1 appeared at 664 nm where the r value is highest when using assumed λ2 and λ3 values of 700 nm and 750 nm, respectively. λ2 and λ3 appeared at 695 nm and 736 nm when using the tuning method. These results showed a linear relationship between the three-band algorithm and Chl-*a* concentration with a R^2 value of 0.46, as shown in Fig. 4.3.

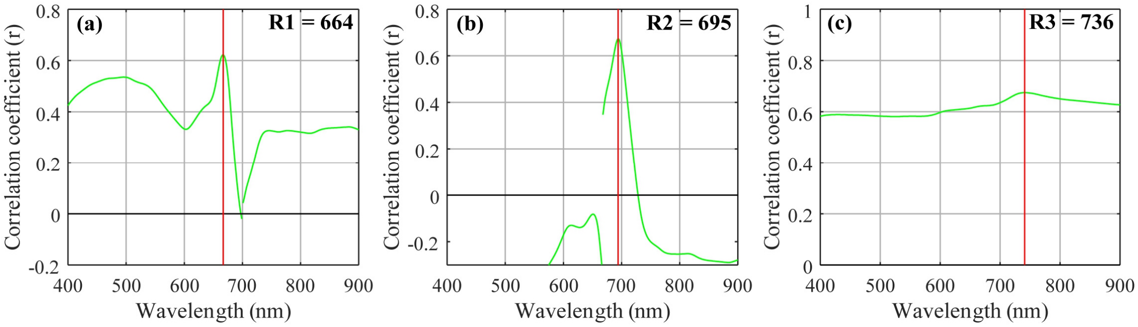

Figure 4.2. Selected wavebands for the three-band model algorithm using the tuning method.

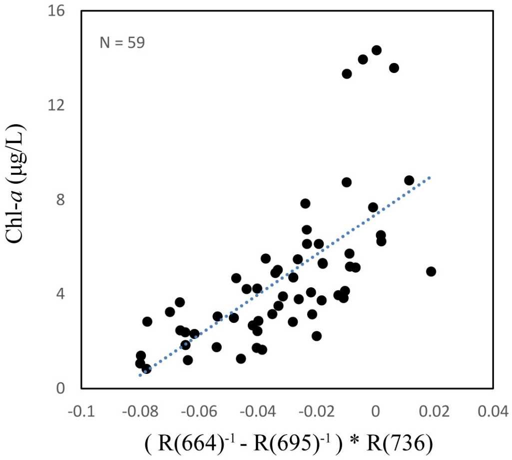

Figure 4.3. Correlation between the three-band model algorithm and measured chlorophyll *a* (Chl-*a*).

The two-band NIR red model results showed an incompact linear relationship between the reflectance ratios of 705 nm and 670 nm and measured Chl-*a*, with a poor R^2 value of 0.17, as shown in Fig. 4.5a. As with three-band model, we tuned the spectral position to obtain the optimal NIR and red wavebands. We initially set the NIR waveband 705 nm and then selected the optimal red region waveband (which we set from 620 nm to 680 nm) based on the highest r value. Fig. 4.4a shows that 666 nm was the optimal red waveband with a r value of 0.43. After fixing the optimal red waveband, we selected the optimal NIR region waveband, which we set from 680 nm to 740 nm. As shown in Fig. 4.4b, we selected 693 nm was selected as the best NIR waveband with a r value of 0.63. Fig. 4.5b shows a linear relationship between the reflectance ratios of 693 nm and 666 nm and measured Chl-*a*, with a R^2 of 0.39.

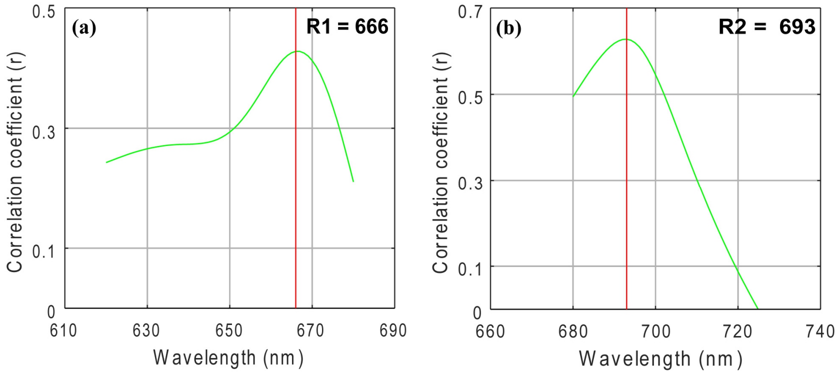

Figure 4.4. Selected wavebands for near-infrared/red algorithm using the tuning method.

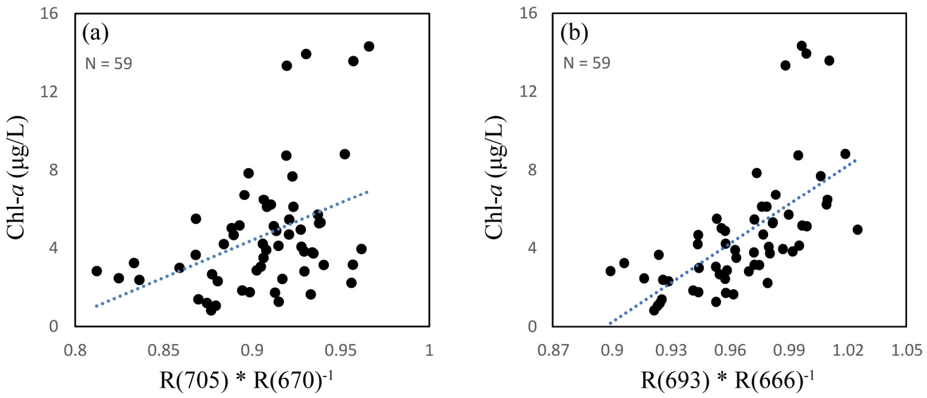

Figure 4.5. Relationship between observed chlorophyll *a* (Chl-*a*) concentration and near-infrared/red reflectance ratio, (a) Ratio of R(705) to R(670); (b) Ratio of R(693) to R(666).

4.3.3. ISE-PLS calibration and validation

Table 4.3 summarizes ISE-PLS calibration and validation results using R_L and FDR for Chl-*a* retrieval. As Table 4.3 shows, ISE-PLS had the same R^2 values (0.83 for both R_L and FDR) and slightly different RMSE values (1.29 for R_L and 1.28 for FDR) for calibration. We also found that ISE-PLS using both datasets had better Chl-*a* retrieval performance than other algorithms, which was indicated by R^2 (0.77 for R_L and 0.78 for FDR) and RPD (2.10 for R_L and 2.13 for FDR) values in the validation results. ISE-PLS using FDR performed marginally better than ISE-PLS using R_L because of the higher R^2 and RPD and lower RMSE (1.47 for R_L and 1.45 for FDR) values for validation. Figures 4.6a and 4.6c show validation plots for ISE-PLS using R_L and FDR, respectively. Both figures show a close linear relationship between predicted and observed Chl-*a* with the exception of a few scatter points.

Because of the iterative stepwise elimination function, we selected the optimal wavebands using ISE-PLS for both R_L and FDR datasets based on the lowest RMSE for validation, as shown in Figs. 4.6b 4.6d. Selected wavebands for R_L ranged from 495 to 496 nm (Perez et al., 2007), 589 to 593 nm (Sasaki et al., 2008), and 660 to 667 nm (Dall'Olmo et al., 2005), which had been proven related to phytoplankton absorption, 544 to 549 nm (Perez et al., 2007), and 689 to 696 nm (Dall'Olmo and Gitelson, 2005), which indicated relationship with Chl-*a* fluorescence, and 730 nm which is also sometimes used for Chl-*a* retrieval (Dall'Olmo et al., 2005). We selected a total of 30 (6%) informative wavebands from all 501 wavebands. And for FDR, we selected 10 (2%) informative wavebands from all 501 wavebands.

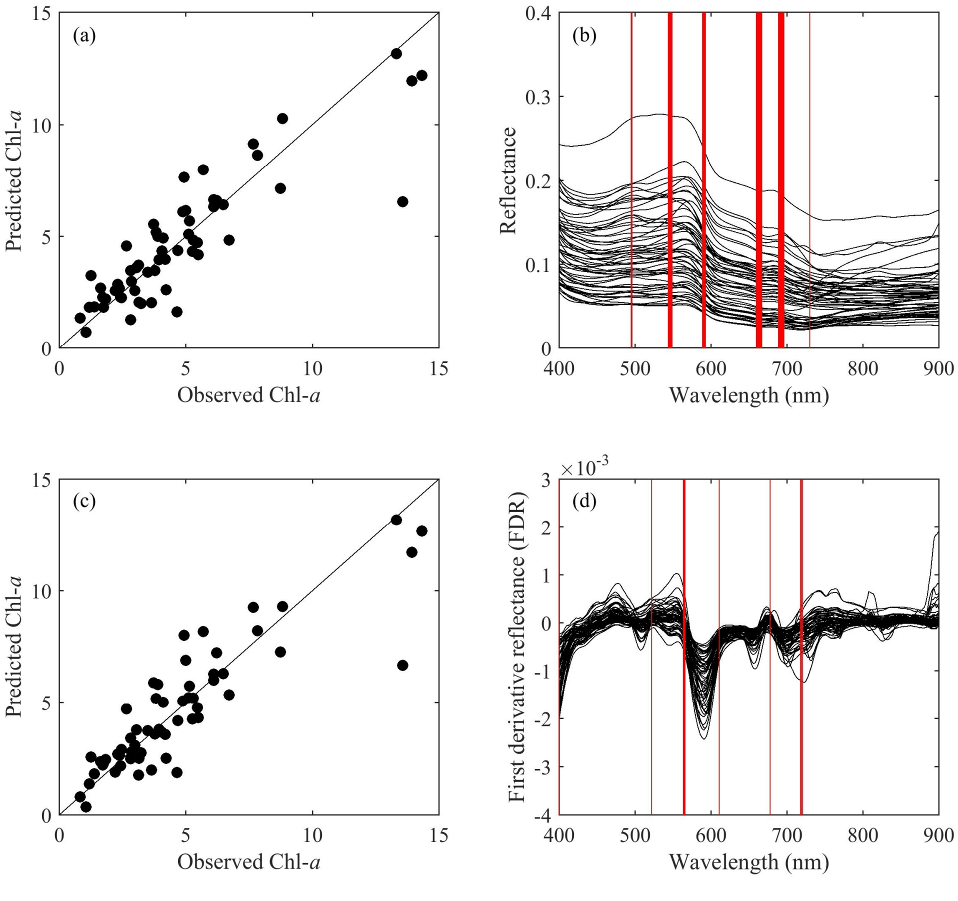

Figure 4.6. Relationship between observed and predicted chlorophyll *a* (Chl-*a*) (a) water-leaving reflectance (R_L); (c) first derivative reflectance (FDR), and selected wavebands

by iterative stepwise elimination partial least squares (ISE-PLS) for Chl-*a* retrieval (b) RL; (d) FDR.

Table 4.3. The coefficient of determination (R^2) and root mean square error (RMSE) for calibration of iterative stepwise elimination partial least squares (ISE-PLS) and leave-one-out (LOO) cross-validation using the entire dataset ($N = 59$), with residual predictive deviation (RPD), number of wavebands, and percent ratio in the full spectrum ($i = 501$).

Dataset	N	Calibration			Validation			Number of selected wavebands	Percentage of selected wavebands (%)
		NLV	R^2	RMSE	R^2	RMSE	RPD		
RL	59	6	0.83	1.29	0.77	1.47	2.1	30	6.0
FDR	59	4	0.83	1.28	0.78	1.45	2.13	10	2.0

N, number of samples; NLV, number of latent variables; RL, water-leaving reflectance; FDR, first derivative reflectance.

4.4. Discussion

4.4.1. Semi-analytical algorithms for each station

Remote sensing methods for retrieving water quality parameters contain spatial and temporal variations because the water body components that affect reflection properties vary in space and time. To further clarify the most fitted Chl-*a* retrieval method in the research area, we analysed algorithms using a separated dataset of six stations. Fig. 4.7 shows the RL and average RL of each station for the research period. As we can see, the average RL of each station shows little difference those of the others. Especially at station 4 (Fig. 4.7d), there is an obvious reflectance peak around 580 nm, which is a result of minimum absorption by all pigments (Yacobi et al., 2011). Fig. 1.2 shows that station 4 is near a river, which could bring various nutrients from land to the coastal area.

Consequently, the highest max Chl-*a* and SD values were obtained at station 4, as shown in Table 4.1.

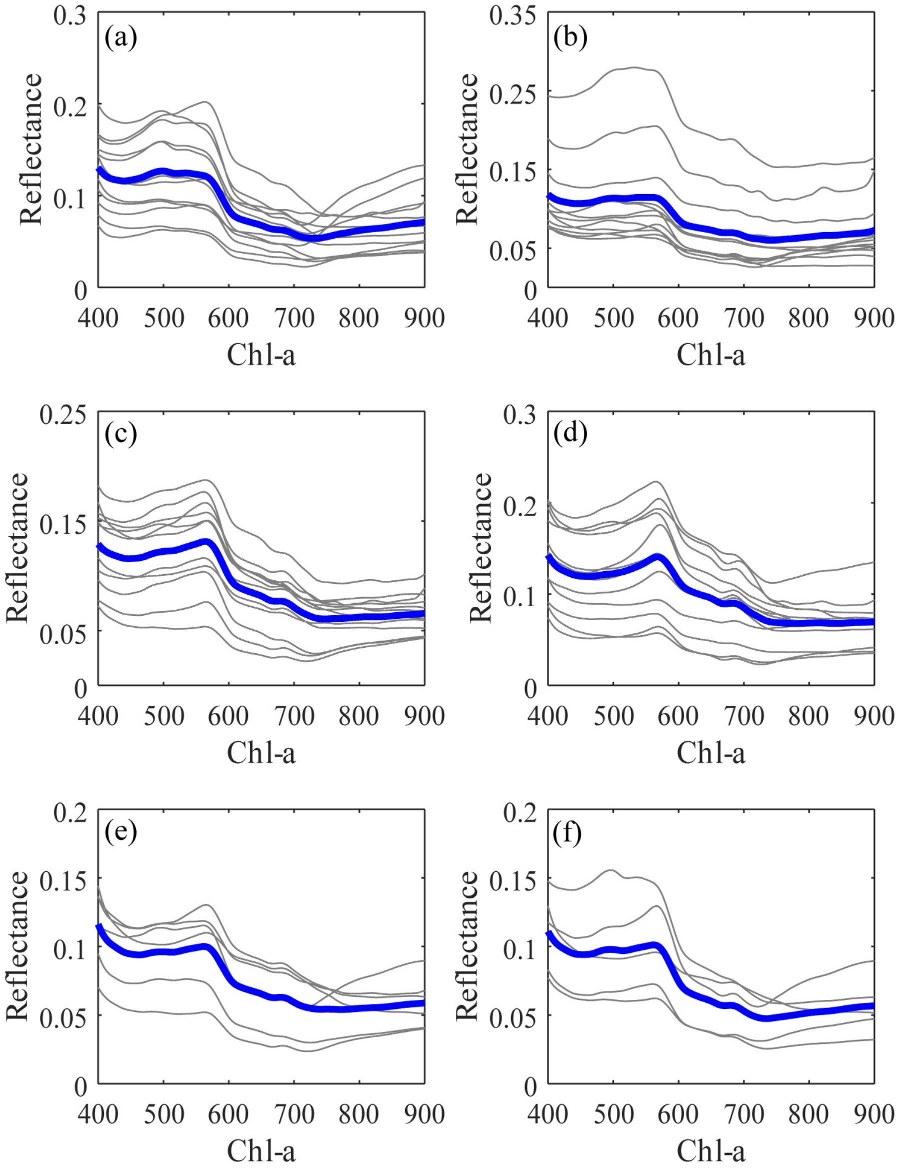

Figure 4.7. Water-leaving reflectance (R_L) spectra with the spectra average (blue line) for each station ((a) to (f) are stations 1 to 6 in turn).

We analysed regressions using all possible band ratios in the 400 to 900 nm range and analysed Chl-*a* concentration for each station, as shown in Fig. 4.8. Two-dimensional correlograms indicate the R^2 distribution for all band ratios (250,000 combinations). The yellow regions indicate high R^2 values for calibration between band ratios and Chl-*a* concentration, with most figures indicating that high R^2 values appear in the NIR and red regions (near 680–710 nm) and green region (near 500–600 nm). However, Fig. 4.8a shows no correlation between NIR/red ratio and Chl-*a* concentration, which may indicate that the NIR/red ratio doesn't fit for water areas with lower and narrower Chl-*a* concentration ranges, as indicated the lowest mean and SD values shown in Table 4.1, which is consistent with a previous study on band ratio analysis (Han and Rundquitst, 1997).

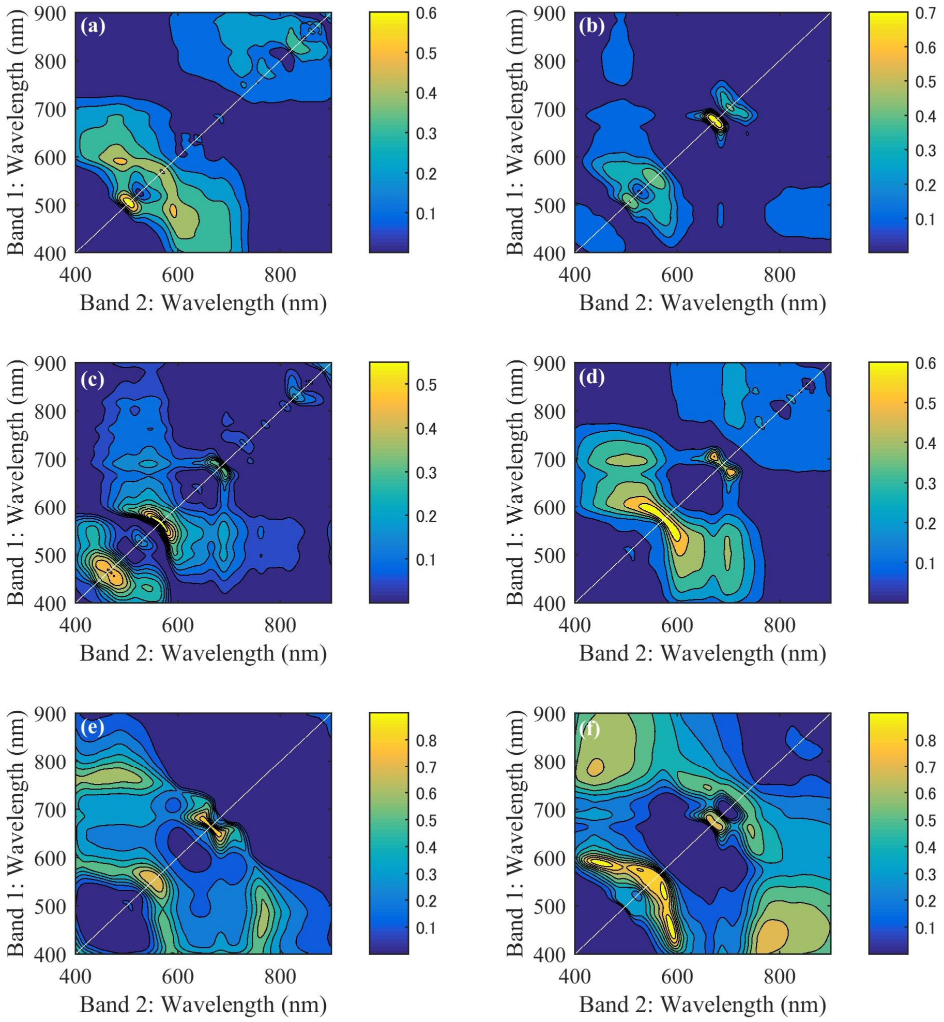

Figure 4.8. Two-dimensional R^2 distributions obtained through sequential regressions using all band ratios and chlorophyll *a* (Chl-*a*) concentrations for each station ((a) to (f) are stations 1 to 6 in turn).

We conducted calibrations between the three-band algorithm and Chl-*a* concentration at each station, the results of which are shown in Fig. 4.9. We selected three optimal wavebands using a tuning method before conducting calibration for each station. It is apparent that station 4 performed better than other stations (1, 2, and 3) with the same dataset number (N = 12), with a R^2 value of 0.66, using wavebands of 674, 705, and 750 nm. However, we obtained a poor R^2 at station 1, which had the lowest Chl-*a* concentration in this study, using wavebands of 664, 689, and 750 nm. These results may indicate that the three-band algorithm performs well in water with relatively higher Chl-*a* concentrations, which is consistent with several previous studies (Dall'Olmo et al., 2003; Song et al., 2013). Figures 4.9e and 4.9f also showed better R^2 values (0.63 at station 5 and 0.81 at station 6) with calibration between the three-band algorithm and Chl-*a* concentration. This provides a possibility of using the three-band algorithm to estimate Chl-*a* in these areas; nevertheless, a shortness of data (N = 6 at station 5, N = 5 at station 6) may also provide uncertainty to the results.

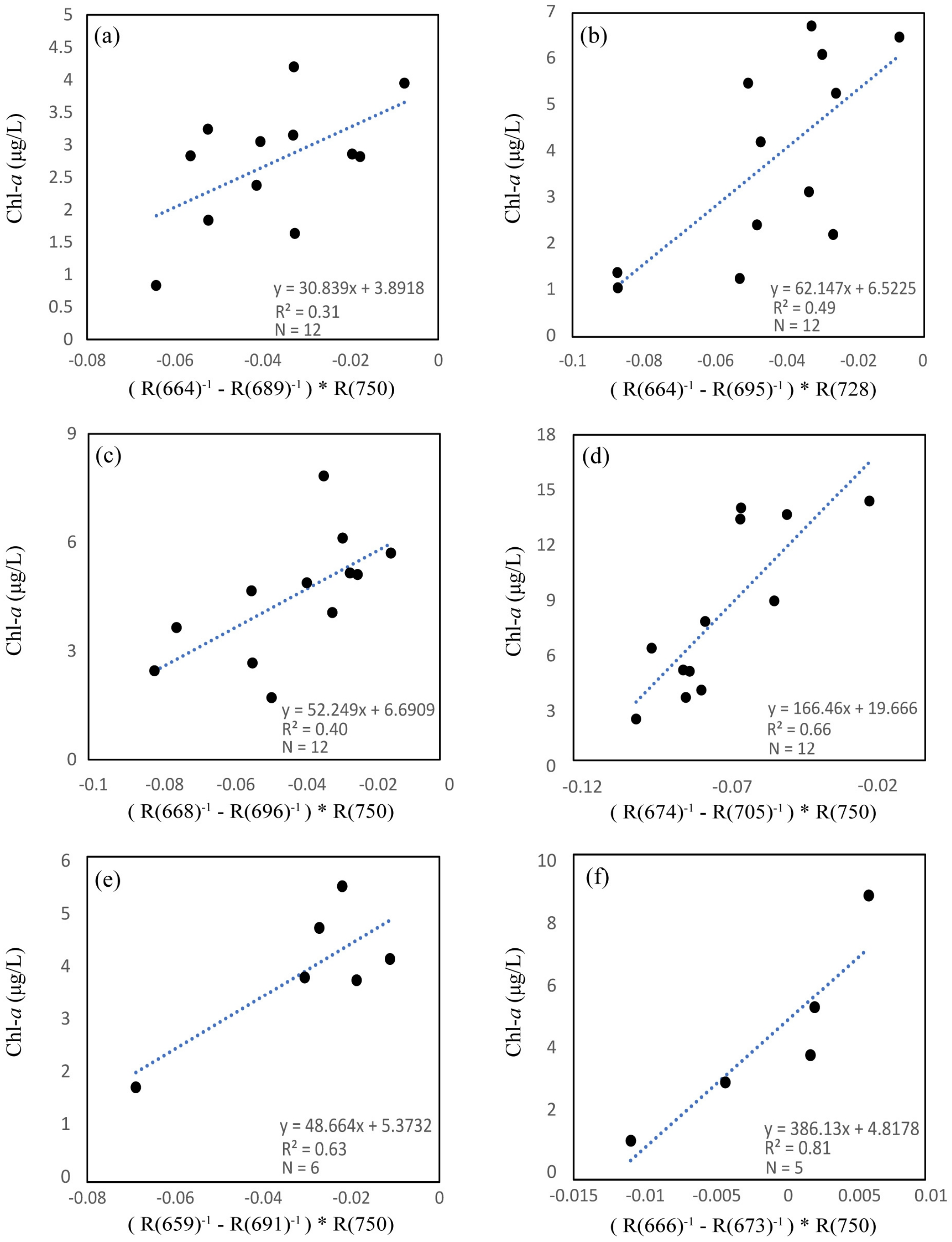

Figure 4.9. Calibrations between the three-band model algorithm and chlorophyll *a* (Chl-*a*) concentrations ((a) to (f) are stations 1 to 6 in turn).

4.4.2. ISE-PLS performance

ISE-PLS using both R_L and FDR performed better than other algorithms, including the OC, three-band model, and NIR/red two-band model algorithms, with all possible band ratios indicated by higher R^2 values. In addition, ISE-PLS using FDR performed better than ISE-PLS using R_L as indicated by higher R^2 and RPD values and lower RMSE for validation. This may have resulted from derivative analysis reducing random noise and removing the effects of suspended matter on Chl-*a* concentration estimates (Song et al., 2013). Because station 4 may have been affected by river nutrients, we carried out ISE-PLS regressions using the R_L dataset except for at station 4 to decrease the impact of different water types. Fig. 4.10 shows the validation plot between observed and predicted Chl-*a*, which was obtained using the LOO method in the ISE-PLS regression. As we can see, the maximum Chl-*a* concentration (from 14.33 to 8.74 μg/L) decreased after removing the station 4 dataset. Results shows a close linear relationship between observed and predicted Chl-*a*; however, compared to the R^2 value (0.77) obtained by ISE-PLS validation using all datasets (N = 59), a relatively lower R^2 value (0.72) was obtained using the datasets except station 4 (N = 47), which may indicate that ISE-PLS performed better in water areas with a wide range of Chl-*a* concentrations.

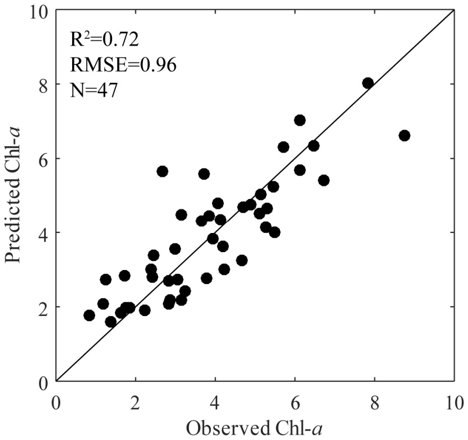

Figure 4.10. Validation of iterative stepwise elimination partial least squares (ISE-PLS) method using lower chlorophyll *a* (Chl-*a*) concentration dataset (station 4 not included).

4.5. Conclusions

In this study, we developed various models for estimating water Chl-*a* concentration in the Seto Inland Sea, including ISE-PLS using both R_L and FDR and other methods such as OC, three-band model, and two-band model algorithms. Our results showed that the ISE-PLS method is effective for predicting Chl-*a* concentration in the Seto Inland Sea using *in situ* measured spectral data. With a higher prediction accuracy, ISE-PLS also selects important wavebands that match previously published studies. Additionally, ISE-PLS using FDR is marginally enhanced compared to using R_L for Chl-*a* retrieval. However, OC algorithms are not robust in this present study, and three-band and two-band model algorithms did not perform well in water areas with lower Chl-*a* concentration. Our results also indicate that the ISE-PLS method can perform better when used in water areas with a wide range of Chl-*a* concentrations. These results provide potential insights into coastal water quality assessment by using a Chl-*a* estimation method with hyperspectral measurements.

Chapter 5: Development of new salinity algorithm in the Seto Inland Sea using simple bio-optical method

5.1 Introduction

The Seto Inland Sea is a semi-enclosed coastal sea in Japan. It is well known as one of the world's most productive treasury of fishery stocks due to the variety of fish species and their abundance (Hashimoto et al., 1997). There are about 35 million people live around this sea, basing on this, the Seto Inland Sea is also one of the most industrialized regions in Japan (Yoshie et al., 2011). The marine ecological environment inevitably affected by human activities. To assessment the sea water quality efficiently, which is important for management fishery industry, several researches have been conducted mainly using chlorophyll-*a* (Chl-*a*) (e.g., Komorita et al., 2016; Nishijima et al., 2016; Kimura et al., 2001). Sea surface salinity (SSS) is one of the key variables for monitoring and modeling ocean circulation (Berger et al., 2002), furthermore, it is one of the prime determinants of the environment in which fish and other marine life live (Marghany, 2010).

Colored dissolved organic matter (CDOM) is an important optical component of marine and fresh water (Bowers & Brett, 2008). The concentration of CDOM can be obtained by a multi-componential based bio-optical model, which depends on the inherent optical properties (IOPs) in local sea water. Sugihara et al. (1985) proposed a bio-optical model in the Tokyo bay which was fitted for the water quality parameters retrieval. Besides, a number of previous researches have shown the CDOM absorption coefficient at 440 nm ($a_y(440)$) is often inversely correlated with salinity with a linear correlation (e.g., McKee et al., 1999; Siddorn et al., 2001; Bowers et al., 2000). This proves that CDOM can be used as a proxy for SSS. However, this finding has not been validated when $a_y(440)$ as a proxy for model building such as bio-optical model by previous researches.

In this study, we made the sea water parameters reversion by combining the bio-optical model and the feature which salinity is linearly correlated with CDOM. Results were validated by estimated $a_y(440)$ and *in situ* SSS.

The objective of this study is to test the applicability of $a_y(440)$ as a proxy for SSS; Furthermore, to establish the model for estimating SSS using *in situ* reflectance and water quality datasets from FY2015 to FY2016.

5.2 Materials and methods

5.2.1 Study area

The research sites locations are shown in Fig. 1.2. Six sites were selected from the center part of Seto Inland Sea near Fukuyama city.

5.2.2. Data acquisition and preprocessing

For *in situ* survey, in total of six times surveys were conducted during December 16, 2015 and December 21, 2016 by using a ship, from which 32 datasets were obtained in six sites except for 4 datasets didn’t be obtained due to the weather condition. Water surface reflectance measurements were carried out using a portable MS-720 spectroradiometer ((EKO Instruments, Tokyo, Japan) with spectral range of 350–1050 nm coupled with data exporting software. The spectral readings were done by approximately 1 m above water surface between 9:00 and 11:00. Chl-*a* and SSS were measured using CTD instrument just under surface water. The statistics for Chl-*a* and SSS are shown in Table 5.1.

Table 5.1. Descriptive statistics of Chl-*a* and SSS

	Chl-*a* (μg/L)	SSS (psu)
Min	1.39	29.48
Max	23.34	32.80
Average	6.74	31.17
N	32	32

For the spectral data, the marginal ranges 350–399 nm and 901–1050 nm from each spectrum were removed due to noise. After that, the spectral data was smoothed using Savitzky-Golay method with 15 smoothing points. The reflectance spectra from all data and an average spectrum are shown in Fig. 5.1.

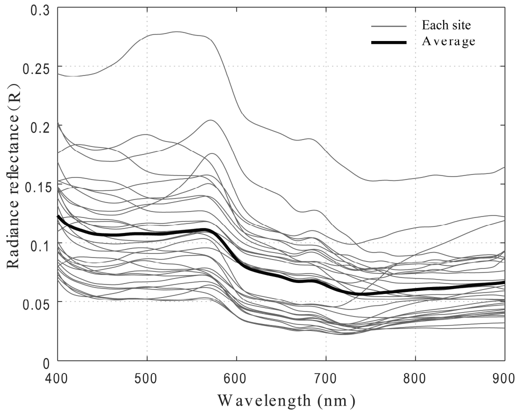

Figure 5.1. Reflectance spectra for all data (*N*=32)

5.2.3. Bio-Optical model

According to Oyama et al. (2009), the radiance reflectance $R(\lambda)$ is represented by the following equation,

$$R(\lambda) = \pi R_{rs}(\lambda) \tag{5.1}$$

where $R_{rs}(\lambda)$ is remote sensing reflectance at the wavelength λ just above the water, which usually obtained by satellite sensor or other instruments is used to predict the water quality parameters. The $R_{rs}(\lambda)$ is also given by

$$R_{rs}(\lambda) = 0.52 * r_{rs}(\lambda)/(1 - 1.7 * r_{rs}(\lambda)) \tag{5.2}$$

$$r_{rs}(\lambda) = 0.33 * b_b(\lambda)/a(\lambda) \tag{5.3}$$

where $b_b(\lambda)$ is the total backscattering coefficient and $a(\lambda)$ is the total absorption coefficient of the sea water. Sugihara et al. (1985) proposed the $a(\lambda)$ equation based on the Tokyo bay sea water characteristics, which can be expressed as below,

$$a(\lambda) = a_w(\lambda) + a_c(\lambda)[Chl\_a] + a_y(\lambda_0)\exp\{-S(\lambda - \lambda_0)\} \tag{5.4}$$

$$b_b(\lambda) = b_w(\lambda)/2 + B_p \tag{5.5}$$

where $a_w(\lambda)$ is the absorption coefficient of pure sea water, $a_c(\lambda)$ is the specific absorption coefficient of Chl-a, $a_y(\lambda_0)$ is the absorption coefficient of CDOM at wavelength λ_0, the λ_0 equal to 440 nm and the spectral slope S has been calculated as 0.014 nm^{-1} (Lahet et al., 2000). $b_w(\lambda)$ is the scattering coefficient of pure sea water and B_p is the backscattering coefficient of particles. Combining the above equations, we can get the relation as follows,

$$R_{rs}(\lambda) = \frac{0.33}{\pi} * \frac{b_w(\lambda)/2 + B_p}{a_w(\lambda) + a_c(\lambda)[Chl\_a] + a_y(\lambda_0)exp\{-S(\lambda - \lambda_0)\}} \tag{5.5}$$

if we introduce the *X* as the coefficients, the following equation can be obtained,

$$X_1[Chl\_a] + X_2[B_p] + X_3[a_y(440)] = Y \tag{5.6}$$

where

$X_1 = R_{rs}(\lambda)a_c(\lambda)/\pi$

$X_2 = -0.33/\pi$

$X_3 = R_{rs}(\lambda)exp\{-0.014(\lambda - 440)\}/\pi$

$Y = 0.33 * b_w(\lambda)2\pi - a_w(\lambda)R_{rs}(\lambda)/\pi$

when we obtain corrected R_{rs} data beyond 3 bands, it will be possible to retrieve the three unknown parameters Chl-*a*, B_p and a_y440 at the same time to solve a simultaneous linear equation. For the equation computation, the $a_w(\lambda)$, $b_w(\lambda)$ and $a_c(\lambda)$ are needed, in this study, the three coefficients were obtained from Pope and Fry (1997), Morel (1974), and Smith and Baker (1987), respectively. The parameters for equation computation are shown in Table 5.2.

Table 5.2. Parameters for equation computation

wavelength (nm)	a_w	a_c	b_w
600	0.2224	0.06	0.000762
680	0.465	0.08	0.000444
700	0.624	0.02	0.000392

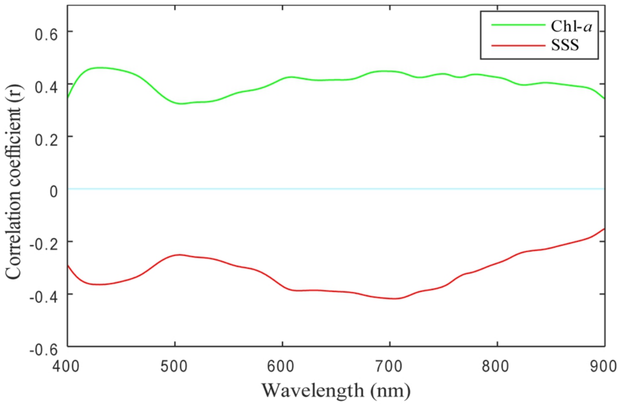

Figure 5.2. Correlation between reflectance and Chl-*a* and SSS

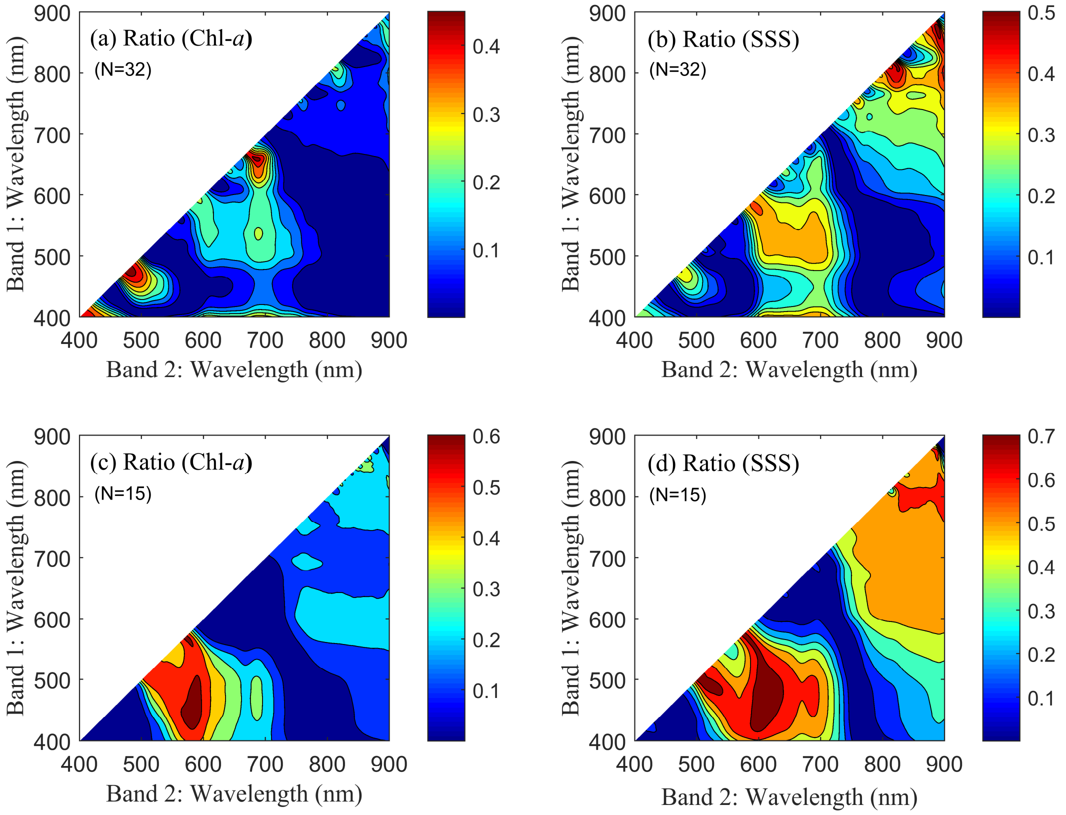

Figure 5.3. R^2 distribution using two wavebands ratio (all datasets (N=32) for Chl-*a* (a) and SSS (b); winter datasets (N=15) for Chl-*a* (c) and SSS (d)).

5.3. Results and discussion

5.3.1. Correlations analysis

In this study, linear correlation analysis was undertaken between water quality parameters and surface reflectance at each wavelength. As we can see from Fig. 5.2, Chl-*a* has a proportional relation with each waveband, while SSS shows inversely relation with wavebands. The absolute value of both correlation coefficient is not satisfactory, but we can still find several higher correlated wavebands (e.g., around 600 nm and 700 nm) both for Chl-*a* and SSS. As a consequence, the reflectance from 600nm and 700nm were selected in the bio-optical model, as shown in Table 5.2.

The algorithm of waveband ratio is widely used for water Chl-*a* retrieval, in this study, we performed the ratio algorithm both for Chl-*a* and SSS, the results are shown in Fig. 5.3. The 2-D correlograms represent the distribution of R^2 which obtained by sequential regression of all possible reflectance ratios in the spectral of 400–900 nm. Results show that the R^2 for winter season (N=15) have an obviously improvement both for Chl-*a* and SSS compared with all year datasets (N=32), the higher value areas (i.e., white spots) are corresponded to red (around 600 nm) and blue (around 450 nm) wavelength both for Chl-*a* and SSS.

5.3.2. Validation

The relationship between *in situ* measured SSS and the estimated $a_y(440)$ are shown in Fig. 5.4. The validation using datasets from winter season were compared with the validation using all datasets. As we can see, estimated $a_y(440)$ has an inversely relationship with *in situ* SSS for both datasets, this finding is in accordance with previous study which conducted by Bowers et al. (2000). However, the R^2 calculated from all datasets is not a satisfactory value (R^2=0.45), which means the predicted $a_y(440)$ may not be suitable for the SSS prediction in all year. It is encouraging that the datasets for winter season has an enhanced R^2 value (R^2=0.66) for the relationship between estimated $a_y(440)$ and *in situ* SSS, the result is in coincident with the correlation results by ratio algorithm, which may indicate that the seasonal variable water components can impact the prediction using bio-optical model. This problem can be solved by the sea water components variation investigation and corrected IOPs selection with regard to different time. Above all, the results indicates the *$a_y(440)$* is potential for SSS prediction using bio-optical model, furthermore, building the predictive model for SSS in Seto Inland Sea.

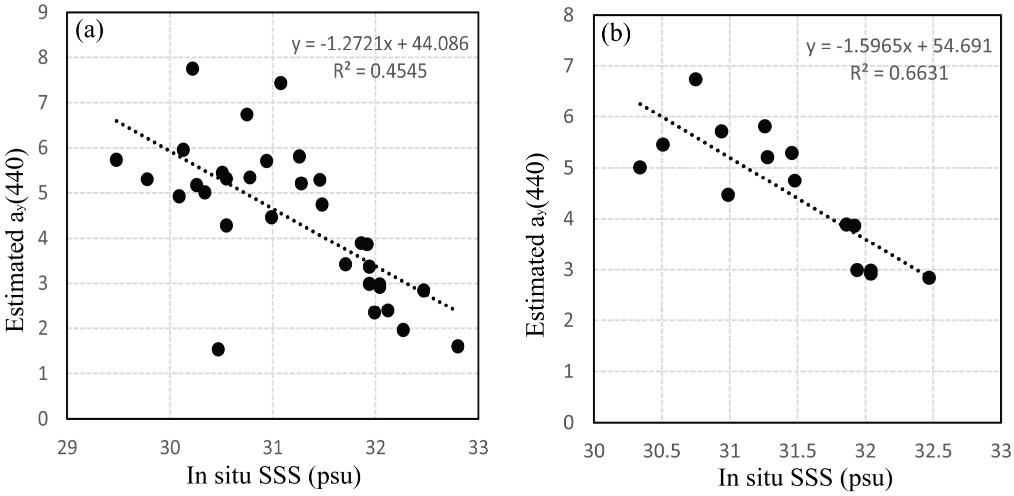

Figure 5.4. Relationship between the measured SSS and the estimated *$a_y(440)$* (using all datasets *N*=32 (a) and datasets from winter *N*=15 (b)).

5.4. Conclusion

The present study developed the model based on the CDOM absorption coefficient in 440 nm for estimating SSS concentration in Seto Inland Sea. Our results showed that the *$a_y(440)$* is a valuable parameter which can be used for modeling to estimate SSS; The bio-optical model can be used in Seto Inland Sea for water quality estimation. These results provide useful insights for future analyses on the assessment of sea water quality, especially when using satellite imagery.

Chapter 6: Validation and modification of new salinity algorithm in the Seto Inland Sea using simple bio-optical method

6.1. Introduction

Coastal area is extremely important for the aquaculture such as oysters, shrimp fisheries, which is helpful for the global food shortage. Japan is one of the world's most important consumers of fishery products, and with a coastline of 29,751 km, Japan is also one of the largest fishery factories. A combination of warm and cold currents flows along the coastal area, creating one of the most abundant fishing grounds in the world (FAO, 2009). Cultivation of lavor is one of the most important aquaculture industries in Japan. However, color fading in recent years has become a problem in various areas such as the Ariake Sea, the Seto Inland Sea, Ise Bay, Mikawa Bay and Tokyo Bay. The main reason of color fading is lack of nutrient salt. Specifically, Dissolved Inorganic Phosphate (DIP) is said to be lacking in Tokyo Bay, Dissolved Inorganic Nitrogen, inorganic nitrogen concentration (DIN) in the Seto Inland Sea and Mikawa Bay is said to be insufficie. Therefore, the observation of DIP and DIN is the most important subject as a survey item for "discoloration" countermeasures. However, it generally takes time and effort to measure nutrient salt, and the on-site type sensor is expensive (several million yen or more), which is unsuitable for planar observation. On the other hand, recent studies have revealed that in the Seto Inland Sea, salinity and DIN have a very high negative correlation between January and February at the age of aquaculture. This is due to the fact that DIN supplied from rivers exists more in low salt waters. In addition, it has been found that colored dissolved organic matter (CDOM) and its associated salt can be detected in a noncontact manner by strong light absorption of water containing organic matter flowing out of a river to ultraviolet to blue (Binding & Bower ,2003; Qing et al, 2013).

In the field of remote sensing, there have been few studies in which coastal nutrients have been contactlessly observed. This is because nutrient salts themselves can't be colored with water even at high concentrations, so it is basically difficult to measure with optical sensors. Salinity estimation by RS technology has recently been launched by satellites such as SMOS (Europe) and Aquarius (USA) which measure oceanic salinity with microwaves at a resolution of about 50 to 100 km, but at the present time, the only technique to measure is to apply the optical method used in the CDOM product of the ocean color sensor. This method utilizes the relationship between light absorption (g 440) and CDOM or "Gelbstoff" and the salt content, and the relationship with reflectance is theoretical by Bowers et al. (2000). In this method, the property of exponentially absorbing the light in the blue to ultraviolet range is utilized as the CDOM concentration increases. As it is known that the concentration of CDOM in terrestrial water is highly negative correlated with salinity in the estuary area, salinity can be estimated by RS technique as a result (Kowalczuk et al. 2006). However, this technology has only been tried in some sea areas such as the Baltic Sea, and case studies are overwhelmingly deficient. In addition, the model is mostly a simple model using two wavelength ratios.

On the other hand, the satellite ocean color sensor, GCOM-C SGLI developed in Japan was launched in December 2017 and is currently checking the data for one year. This satellite sensor has a 250 m resolution capable of estimating the water quality of more complex coastal areas where there are lavor farms and has the ability to monitor coastal water quality (Chl-*a* and SS) in the observation cycle of 2-3 days have. However, salinity products have not been studied yet. Therefore, in this study, the salinity estimation method was examined by using a linear Bio-optical model, which is somewhat complicated than the non-arithmetic algorithm, for salinity estimation using the SGLI band immediately before SGLI stationary operation.

6.2. Materials and Methods

6.2.1. Overview of GCOM-C/SGLI

The SGLI is the name of "multi-wavelength optical radiometer" mounted on the GCOM-C satellite. SGLI is a successor sensor of the global imager (GLI) mounted on ADEOS II, and performs multiband observation from near ultraviolet to thermal infrared region (380 nm to 12 μm). The observation width is 1150 km, the quantization is 12 bits, the observation period is 2 days at mid latitude (around 35 °), and it is within 3 days even at low latitude. The main characteristic of SGLI as an ocean observation satellite is high spatial resolution compared with the conventional ocean color sensor of 250 m spatial resolution. The observation band of SGLI (only the non-polarized part) is shown in Table 6.1.In this study, six visible bands of VN 2 to VN 7, which can be compared with effective spectroscopic measurement data, were targeted. The selected wavelength range almost agrees with the MODIS sensor of 1 km resolution.

Table 6.1. Specification of GCOM-C/SGLI

Channel	Band center (nm)	Band width (nm)	IFOV (m)
VN1	380	10	250
VN2	412	10	250
VN3	443	10	250
VN4	490	10	250
VN5	530	20	250
VN6	565	20	250
VN7/8	674	20	250
VN9	763	12	250/1000
VN10/11	869	20	250

6.2.2. Salinity estimation algorithm for coastal waters

6.2.2.1. Two-band model

Salinity estimation algorithm of coastal area by optical remote sensing is generally estimated through absorption coefficient of dissolved organic substance Gelbstoff (also called yellow substance, Gilvin, CDOM etc) contained in water. Specifically, it is known that the absorption coefficient (g_{440}) of Gelbstoff at 440 nm in coastal water has a high correlation with salinity (*Sal*) (Bowers, 2000; Binding & Bower ,2003).

$$g_{440} \propto Sal \tag{6.1}$$

Therefore, the salinity can be estimated if the g_{440} can be estimated from the reflectance observed by the satellite.

On the other hand, the remote sensing reflection factor R - right under the water surface is represented by the following model.

$$R^{-}(\lambda) = \frac{L_u(\lambda)}{E_d(\lambda)} \propto \frac{b_b(\lambda)}{a(\lambda)} \tag{6.2}$$

where L_u is the upward radiance, Ed is the downward irradiance, a is the total absorption coefficient, b is the total scattering coefficient, and γ is a constant. Of these, a and b are further formulated as follows.

$$a(\lambda) = a_w(\lambda) + a_{ps}(\lambda)[Chl\_a] + a_{ms}(\lambda)[MSS] + a_{ys}(\lambda)[g_{440}] \tag{6.3}$$

$$a_{ys}(\lambda) = \exp(-S(\lambda - \lambda_0)) \tag{6.4}$$

$$b_b(\lambda) = b_{bw}(\lambda) + b_{bps}(\lambda)[Chl\_a] + b_{bpm}(\lambda)[MSS] \tag{6.5}$$

where a_w is the absorption coefficient of pure water, and a_{ps}, a_{ms}, and a_{ys} are the specific absorption coefficients of [*Chl_a*], [*MSS*], and [g_{440}], respectively. Also, S is a constant, λ is a wavelength, λ_0 is a fundamental wavelength (often 440 nm is used for salinity estimation). Furthermore, b_w is the scattering coefficient of pure water, b_{bps}, b_{bms} are the specific scattering coefficients of [*Chl_a*] and [*MSS*], respectively. If the scattering coefficient b is much smaller than a and the following relationship holds, the reflectance ratio of two wavelengths (1, 2) is expressed as follows.

$$\frac{R_1}{R_2} = \gamma\left(\frac{a_2}{a_1}\right) = \gamma\left(\frac{a_{w2} + a_{ps2}[Chl\_a] + a_{ys2}[g_{440}]}{a_{w1} + a_{ps1}[Chl\_a] + a_{ys1}[g_{440}]}\right) \tag{6.6}$$

When you select a red band (red) without absorption of yellow substance for wavelength 1, a_{ys}=0, so equation (6.6) is changed to the following equation.

$$\frac{R_{red}}{R_2} = \gamma\left(\frac{a_{w2} + a_{ps2}[Chl\_a] + a_{ys2}[g_{440}]}{a_{w(red)} + a_{ps(red)}[Chl\_a]}\right) \tag{6.7}$$

Solving for g_{440} used for salinity estimation using equation (6.7) yields the following equation:

$$g_{440} = \left(\frac{a_{w(red)} + a_{ps(red)}[Chl\_a]}{a_{ys2}\gamma}\right)\left(\frac{R_{red}}{R_2}\right) - \frac{a_{w2} + a_{ps2}}{a_{ys2}} \tag{6.8}$$

In equation (6.8), assuming that a_w (red) is significantly larger than a_{ps} (red) and assuming that a_{ps2} is significantly larger than a_{w2}, it can be expressed by the following equation.

$$g_{440} = \left(\frac{a_{w(red)}}{a_{y2}\gamma}\right)\left(\frac{R_{red}}{R_2}\right) - \left(\frac{a_{ps2}}{a_{ys2}}\right) \tag{6.9}$$

Finally, we developed a simple regression model that estimates salinity from reflectance ratios of the following two wavelengths using equation (6.1) and equation (6.9) as follows.

$$Sal = \alpha\left(\frac{R_{red}}{R_2}\right) + \beta \tag{6.10}$$

here, α and β are regression constants obtained by regression. In addition, red band of wavelength 1 is around 665 to 670 nm and wavelength 2 is around 412 to 555 nm (Bowers, 2000; Binding & Bower ,2003).

6.2.2.2. Linear bio-optical model

Apart from the above-described two-band model, a technique for obtaining *Chla* and g_{440} from multiple wavelengths called Linear Matrix Inversion Method (MIM) has been tried in various places although it is not for salinity estimation. Here we consider a model that estimates salinity by applying this MIM developed by Sugihara et al. (1983) in Tokyo Bay, Japan.

$$R_E^{-}(\lambda) = \frac{E_u(\lambda)}{E_d(\lambda)} = 0.33\frac{b_b(\lambda)}{a(\lambda)} \tag{6.11}$$

$$a(\lambda) = a_w(\lambda) + a_{ps}(\lambda)[Chl\_a] + a_{ys}(\lambda)[g_{380}] \tag{6.12}$$

$$a_{ys}(\lambda) = \exp(-0.0167(\lambda - 380)) \tag{6.13}$$

$$b_b(\lambda) = \frac{b_{bw}(\lambda)}{2} + [b_{bp}] \tag{6.14}$$

$$R_E{}^{-}(\lambda) = 0.33 \frac{0.5b_{bw}(\lambda) + [b_{bp}]}{a_w(\lambda) + [Chl\_a]a_{ps}(\lambda) + [g_{380}]exp\{-0.0167(\lambda - 380)\}} \tag{6.15}$$

Further, the equation (6.15) is transformed into the form of a linear equation as follows.

$$X_1[Chl\_a] + X_2[b_{bp}] + X_3[g_{380}] = Y \tag{6.16}$$

$$X_1 = R_E{}^{-}(\lambda)a_{ps}(\lambda) \tag{6.17}$$

$$X_2 = -0.33 \tag{6.18}$$

$$X_3 = 0.33/2R_E{}^{-}(\lambda) \tag{6.19}$$

Since the unknowns are three parameters of $[Chl\_a]$, $[b_p]$, and $[g_{380}]$, if $R_E{}^{-}$ of three or more wavelengths can be obtained, from the equations (16) to (19), using the least squares method. Assuming that there is a relationship such as equation (1) between g380 and salinity obtained in this manner, salinity can be estimated from the reflectance of multiple bands of three bands or more finally.

$$Sal \propto g_{380} \tag{6.20}$$

However, in this case, a_w, a_{ps}, b_w are used as shown in Table 6.2, of which a_w is the value of Pope and Fry (1997) and a_{ps} is the expression of Lee et al. (1998) (This time it was fixed as the average Chl-*a* as 5 mg/m^3), b_{bw} is calculated by Morel (1974).

Table 6.2. IOPs for SGLI band

wavelength (nm)	412	443	490	530	565	674
a_w	0.0047	0.0072	0.015	0.0434	0.0657	0.4494
a_{ps}	0.0687	0.068	0.0463	0.0305	0.0174	0.0326
b_w	0.0039	0.0028	0.0018	0.0013	0.001	0.0005

6.2.3. R_{rs} / Chl-*a* data set

During the period from May 13, 2014 to October 14, 2016, at the 6 stations (Stn.1 to Stn.6) of the Seto Inland Sea off Fukuyama in Hiroshima Prefecture shown in Fig. 1.2, Chl-*a* salt and spectral reflectance of December 16, 2015, January 13, 2016, March 22, June 14, October 13, December 21) were carried out. The Chl-*a* used for the calculation, the salinity is the data of the water depth 0 m. For the measurement, Cyclops-7 manufactured by Turner designs and DS-5 manufactured by Hydrolab were used, respectively. For spectral reflectance measurement, a portable spectroradiometer MS 720 (EKO Ltd.) was used. The MS 720 has the capability to measure the irradiance (Wm^{-2} μm) at a wavelength of 350 to 1050 nm at a step of 3.3 nm. The spectral irradiance on the water surface directly above the water surface and the spectral irradiance of the solar radiation reflected from the white plate Labsphere were measured three times and averaged. The obtained data was converted to irradiance data of 1 nm step using the attached software (MS 720_x 64.Exe). Remote sensing reflectance R_{rs} (sr^{-1}) was calculated by referring to the method of Oyama et al.(2009), Calculating the ratio of the average values and then dividing by the pi.

In addition, due to factors such as the rapid tidal current and the structure of the ship, we could not measure the spectral reflectance considering the positional relationship of ships, instruments, and sun to mitigate the influence of surface reflection.

Therefore, after-processing to alleviate the influence of water reflection light was carried out. As a simple glint correction process of coastal water, there is a method of Kutser et al (2013). Basically, this method subtracts the baseline corresponding to the glint on the assumption that the reflectance of ultraviolet (350-380 nm) and near infrared

(890-900 nm) is 0. However, the R_{rs} data of this time has large unexplained noise at ultraviolet and near infrared wavelengths, and when this method is adopted, a lot of data are calculated as negative. Therefore, this time, only a simple offset correction with 0 as the minimum value of 700 to 800 nm in each data was performed. Fig. 6.1 shows the spectral reflectance characteristics of R_{rs} before and after offset correction.

On the other hand, since the offset corrected spectral reflectance is R_{rs} above the water surface, it can't be compared with the values of the equations (6.2) and (6.11), which are reflectance immediately below the water surface calculated by the model. Therefore, in this study, we compared each data by dividing the expression of Lee et al. (1999) which formulated as equations (6.21) and (6.22).

$$R^{-} = \frac{L_u}{E_d} = \frac{R_{rs}}{0.52 + 1.7R_{rs}} \tag{6.21}$$

$$R_E{}^{-} = \frac{E_u}{E_d} = R^{-}/\pi \tag{6.22}$$

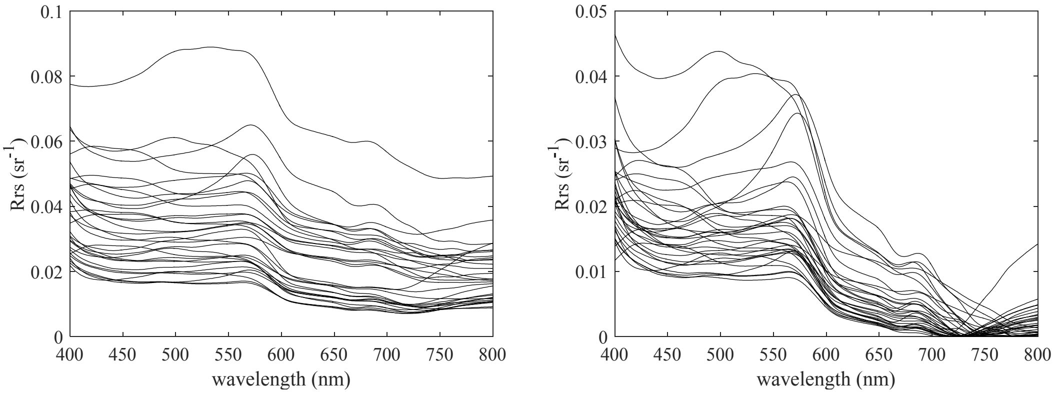

Figure 6.1. Characteristic of measured spectral reflectance. (a) Before the offset correction, (b) After the offset correction.

6.3. Results and Discussion

6.3.1. Two-band model

Fig. 6.2 shows the SSS estimation result by the two-band model shown by the equation (6.10). There was no significant correlation between Type A and Type B. However, a completely different relationship was shown between Type A and Type B data areas (open circles in the figure). This was thought to be due to the difference in the spectral characteristics of the R_{rs} data of both. Therefore, it seems to be effective to analyze the type of both (Type A and Type B) separately. However, unless both types are distinguished automatically, it is not a practical body. When spectral characteristics of both were visually confirmed, the shapes at 412 nm, 443 nm, and 490 nm differed between Type A and Type B. For this reason, as shown in Fig. 6.3, the slopes of these two wavelengths were compared. As a result, two different type of dataset are obviously separated in Fig. 6.3a. This result shows that Type A and Type B can be separated using a scatter diagram of "Slope of R_{rs} at 412 nm and 443 nm" and "Slope at 443 nm and 490 nm".

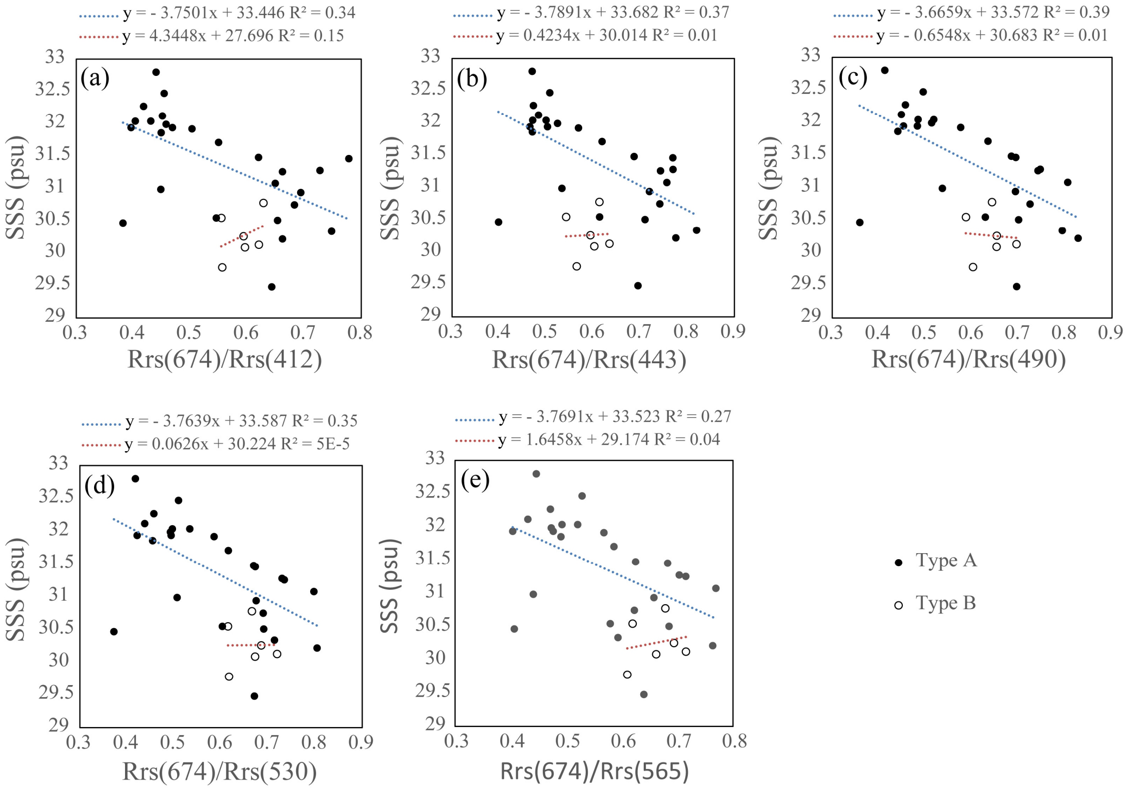

Figure 6.2. Relationship between two band ratio model and SSS. (a) $R_{rs}(674)/R_{rs}(412)$ vs SSS, (b) $R_{rs}(674)/R_{rs}(443)$ vs SSS, (c) $R_{rs}(674)/R_{rs}(490)$ vs SSS, (d) $R_{rs}(674)/R_{rs}(443)$ vs SSS, (e) $R_{rs}(674)/R_{rs}(565)$ vs SSS

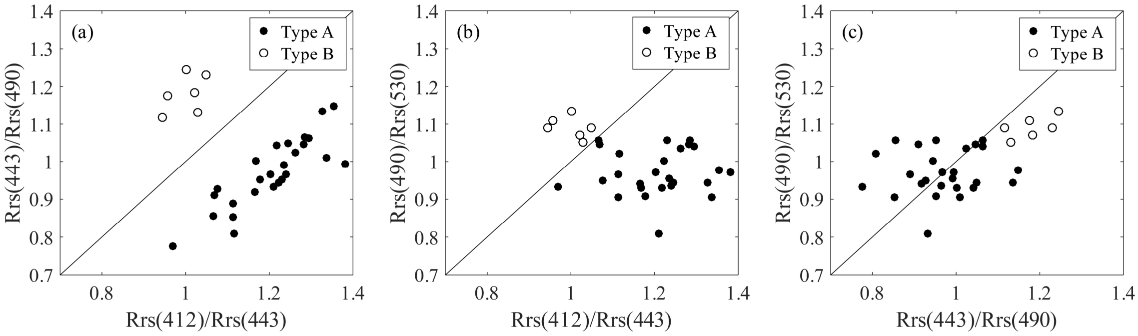

Figure 6.3. Scatter diagram of slope for separation between Type A and Type B. (a) $R_{rs}(412)/R_{rs}(443)$ vs $R_{rs}(443)/R_{rs}(490)$, (b) $R_{rs}(412)/R_{rs}(443)$ vs $R_{rs}(490)/R_{rs}(530)$, (c) $R_{rs}(443)/R_{rs}(490)$ vs $R_{rs}(490)/R_{rs}(530)$.

6.3.2. Linear Bio-optical Model

The relationship between estimated Chl-*a* from bio-optical model and observed Chl-*a* is shown in Fig. 6.4a, the estimated Chl-*a* has a linear relationship with observed Chl-*a* both for Type A and Type B dataset (R^2 = 0.69 for Type A and R^2 = 0.76 for Type B). This result indicate Chl-*a* can be estimated by the bio-optical model with SGLI bands. Fig. 6.4b shows the relationship between estimated g380 and observed SSS, the estimated g380 also shows linear relationship and observed SSS for both type of dataset (R^2 = 0.43 for Type A and R^2 = 0.64 for Type B), which shows the potential of bio-optical model with SGLI bands to estimate SSS in the Seto Inland Sea.

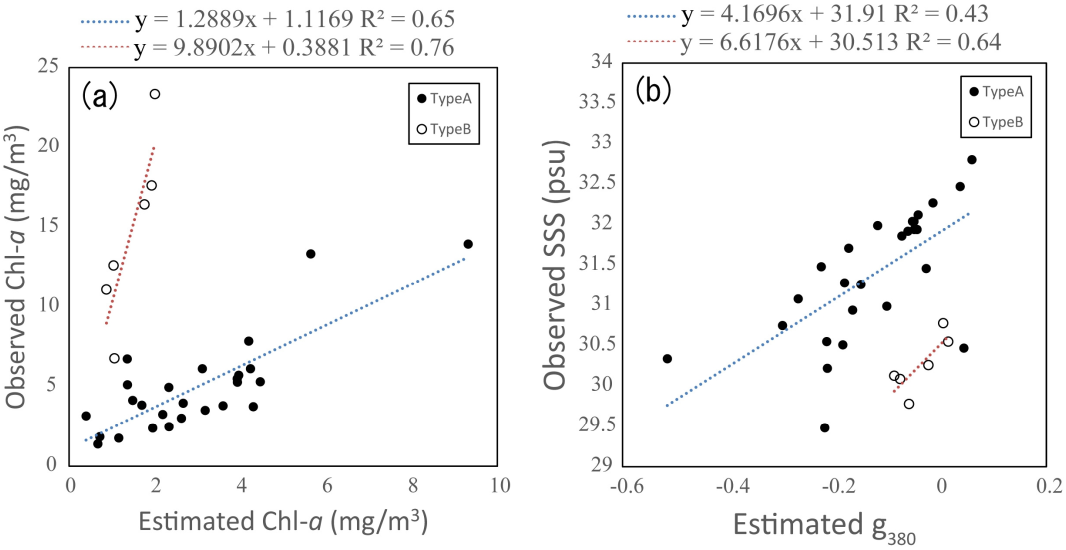

Figure 6.4. Correlation between Estimated value and Observed value under Type A and Type B conditions. (a) Estimated Chl-*a* and observed Chl-*a*, (b) Estimated g_{380} and observed SSS.

However, the black dots and circles from Fig. 6.4b are more incompact than that of Fig. 6.4a, which indicate the accuracy for SSS estimation using bio-optical model with

SGLI bands is no better than Chl-*a* estimation. In addition, the accuracy for SSS estimation for Type A dataset is no better than Type B dataset.

The separation of different type dataset is a useful way to discuss the estimation for parameters (i.e., Chl-*a* and SSS) using complex spectral data. Because spectral reflectance can be affect by different water components, which makes difference for the spectral pattern for different regions. To better understand the relationship between modeled Chl-*a*/SSS and observed Chl-*a*/SSS, Type A and Type B dataset were combined into one figure, as shown in Fig. 6.5. Black dots and circles are close to the 1:1 line, which displays the potential predictive ability both for Chl-*a* and SSS. However, more compact black dots and circles reveal in Fig. 6.5a, indicates the model is more robust for Chl-*a* than SSS.

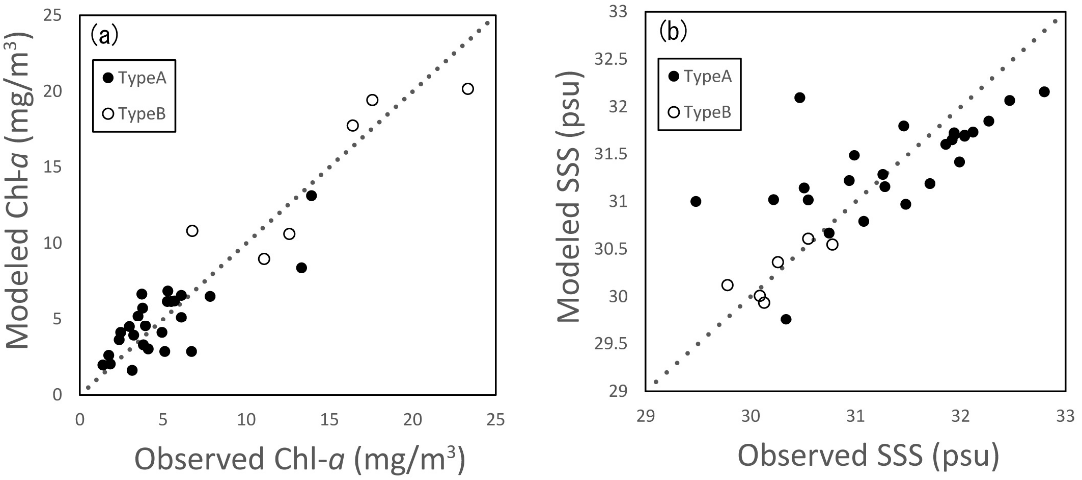

Figure 6.5. Scatter diagram of observed and modeled water quality parameters under Type A and Type B conditons. (a) Estimated Chl-*a* and observed Chl-*a*, (b) Estimated g_{380} and observed SSS.

6.4. Conclusion

In this study, algorithms for SSS estimation with SGLI bands in the Seto Inland Sea were developed, including a two-band ratio and a bio-optical model. Results showed the two-band ratio is not a better method for SSS estimation using SGLI bands with a poor coefficient of determination. On the contrary, proposed bio-optical model based on SGLI bands has a potential predictive ability for SSS estimation in the Seto Inland Sea. In addition, the separation method for spectral patterns can be applied for improving the accuracy.

Chapter 7: Conclusions

The present study develops models for estimating Chl-*a* and TSS concentrations in irrigation ponds using water surface reflectance spectral data. Results showed that PLS regression analysis has high potential for predicting Chl-*a* and TSS based on field hyperspectral measurements, and that ISE wavebands selection in combination with PLS regression analysis can enhance predictive ability. Chl-*a* and TSS concentrations were estimated with high accuracy by using ISE-PLS, which explains 98% of the variance for Chl-*a* and 97% of the variance for TSS. The important wavebands for estimating Chl-*a* and TSS using ISE–PLS represented 16.97% and 8.38%, respectively, of all 501 wavebands over the 400–900 nm range. The selected wavebands approximately match the absorption peaks published by previous researchers. Compared to the estimation of water quality parameters by satellite sensors such as MODIS, ISE–PLS selected more informative wavebands, especially the wavelength at approximately 700 nm. These results provide useful insights for future analyses on the assessment of water quality in irrigation ponds, especially when using satellite imagery.

The author also developed various models for estimating water Chl-*a* concentration in the Seto Inland Sea, including ISE-PLS using both R_L and FDR and other methods such as OC, three-band model, and two-band model algorithms. Results showed that the ISE-PLS method is effective for predicting Chl-*a* concentration in the Seto Inland Sea using *in situ* measured spectral data. With a higher prediction accuracy, ISE-PLS also selects important wavebands that match previously published studies. Additionally, ISE-PLS using FDR is marginally enhanced compared to using R_L for Chl-*a* retrieval.

However, OC algorithms are not robust in this present study, and three-band and two-band model algorithms did not perform well in water areas with lower Chl-*a* concentration. Our results also indicate that the ISE-PLS method can perform better when used in water areas with a wide range of Chl-*a* concentrations. These results provide potential insights into coastal water quality assessment by using a Chl-*a* estimation method with hyperspectral measurements.

The present study developed the model based on the CDOM absorption coefficient in 440 nm for estimating SSS concentration in Seto Inland Sea. Results showed that the $a_y(440)$ is a valuable parameter which can be used for modeling to estimate SSS; The bio-optical model can be used in Seto Inland Sea for water quality estimation. These results provide useful insights for future analyses on the assessment of sea water quality, especially when using satellite imagery.

In addition, algorithms for SSS estimation with SGLI bands in the Seto Inland Sea were developed, including a two-band ratio and a bio-optical model. Results showed the two-band ratio is not a better method for SSS estimation using SGLI bands with a poor coefficient of determination. On the contrary, proposed bio-optical model based on SGLI bands has a potential predictive ability for SSS estimation in the Seto Inland Sea. Besides, the separation method for spectral patterns can be applied for improving the accuracy.

Acknowledgements

I would like to acknowledge to my supervisor Associate Professor Yuji Sakuno of the Graduate School of Engineering, Hiroshima University, Japan, for his kindly encouragement, suggestions, patient guidance and support during my Ph.D. study.

My appreciation is extended to my sub-supervisors, Professor Hidetsugu Iwashita, Professor Yasuaki Doi, and Associate Professor Hidemi Mutsuda, for their constructive comments and suggestions regarding my study.

Then I wish to express my deep appreciation to Associate Professor Kensuke Kawamura for his comments, guidance and energetic support for my master’s research.

I would like to express my deep appreciation to Professor Kazuhiko Koike of the Graduate School of Biosphere Science, Hiroshima University, Japan, for his support and guidance for my field survey. And I appreciate Shizuka Ohara and other Koike’s lab members for their help and support for data collection for my research.

My appreciation is extended to Assistant Professor Naokazu Taniguchi, for his kindly guidance for my doctoral dissertation.

Also I would like to give much thanks to all the members in my lab, as well as many friends in different fields, for their kindly supports, advices, and friendship.

Finally, I would like to appreciate my parents and sister, they give me much encouragement, care, and especially the huge support for pursuing my dream. And my appreciation is extended to Jie Song for her understanding, support, and companion.

References

[1] Abe, H., & Shinohara, S. (1996). A study on irrigation ponds in Higashihiroshima: A statistical approach. *J. Fac. Appl. Biol. Sci. Hiroshima Univ,* 35, pp. 27–34.

[2] Ahn, Y. H., Shanmugam, P., Moon, J. E., & Ryu, J. H. (2008). Satellite remote sensing of a low–salinity water plume in the East China Sea. *Annales Geophysicae,* 26, 2019–2035.

[3] Austin, R. W. (1980). Gulf of Mexico, ocean–color surface–truth measurements. *Boundary–Layer Meteorology,* 18(3), pp. 269–285.

[4] Bash, J., Berman, C., & Bolton, S. (2001). *Effects of Turbidity and Suspended Solids on Salmonids.* Seattle: University of Washington.

[5] Bennet, A., & Bogorad, L. (1973). Complementary chromatic adaptation in a filamentous blue–green alga. *J. Cell Biol.,* 58(2), 419–435.

[6] Berger, M., Camps, A., Font, J., Kerr, Y., Miller, J., Johannessen, J., . . . Attema, E. (2002). Measuring Ocean Salinity with ESA's SMOS Mission – Advancing the Science. *ESA Bulletin – European Space Agency,* 111, 113–121.

[7] Binding, C. E., Bowers, D. G. (2003). Measuring the salinity of the Clyde Sea from remotely sensed ocean color. *Estuarine, Coastal and Shelf Science* 57, 605–611.

[8] Boggia, R., Forina, M., Fossa, P., & Mosti, L. (1997). Chemometric Study and Validation Strategies in the Structure–Activity Relationships of New Cardiotonic Agents. *Quantitative Structure–Activity Relationships,* 16(3), 201–213.

[9] Bowers, D. G., Harker, G. E., Smith, P. S., & Tett, P. (2000). Optical Properties of a Region of Freshwater Influence (The Clyde Sea). *Estuarine, Coastal and Shelf Science,* 50(5), 717–726.

[10] Bowers, D. G., & Brett, H. L. (2008). The relationship between CDOM and salinity in estuaries: An analytical and graphical solution. *Journal of Marine Systems*, 73(1–2), 1–7.

[11] 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. *Journal of Geophysical Research: Oceans,* 103(C13), pp. 31033–31044.

[12] Chang, C., Laird, D. A. (2002). Near–infrared reflectance spectroscopic analysis of soil C and N. *Soil Science*, 167(2), 110–116.

[13] Chen, D., Cai, W., & Shao, X. (2007). Representative subset selection in modified iterative predictor weighting (mIPW) — PLS models for parsimonious multivariate calibration. *Chemometrics and Intelligent Laboratory Systems,* 87(2), 312–318.

[14] D'Acqui, L. P., Pucci, A., & Janik, L. J. (2010). Soil properties prediction of western Mediterranean islands with similar climatic environments by means of mid–infrared diffuse

reflectance spectroscopy. *Soil Science,* 61(6), 865–876.

[15] D'Archivioa, A. A., Maggi, M. A., & Ruggieri, F. (2014). Modelling of UPLC behaviour of acylcarnitines by quantitative structure–retention relationships. *Journal of Pharmaceutical and Biomedical Analysis,* 96, pp. 224–230.

[16] Dall'Olmo, G., & Gitelson, A. A. (2005). Effect of bio–optical parameter variability on the remote estimation of chlorophyll–a concentration in turbid productive waters: experimental results. *Applied Optics,* 44(3), 412–422.

[17] Dall'Olmo, G., Gitelson, A. A., & Rundquist, D. C. (2003). Towards a unified approach for remote estimation. *Geophysical Research Letters,* 30(18).

[18] Dall'Olmo, G., Gitelson, A. A., Rundquist, D. C., Leavitt, B., Barrow, T., Holz, J. C. (2005). Assessing the potential of SeaWiFS and MODIS for estimating chlorophyll concentration in turbid productive waters using red and near–infrared bands. *Remote Sensing of Environment*, 96, 176−187.

[19] Davies–Colley, R. J., & Smith, D. G. (2007). TURBIDITY SUSPENI)ED SEDIMENT, AND WATER CLARITY: A REVIEW. *J. Am. Water Resour. Assoc.,* 37(5), 1085–1101.

[20] Derbalah, A. S., Nakatani, N., & Sakugawa, H. (2003). Distribution, seasonal pattern, flux and contamination source of pesticides and nonylphenol residues in Kurose River water, Higashi–Hiroshima, Japan. *Geochem. J.,* 37(2), 217–232.

[21] El–Serehy, H. A., Abdallah, H. S., Al–Misned, F. A., Al–Farraj, S. A., & Al–Rasheid, K. A. (2018). Assessing water quality and classifying trophic status for scientifically based managing the water resources of the Lake Timsah, the lake with salinity stratification along the Suez Canal. *Saudi Journal of Biological Sciences*.

[22] FAO. (2009). *Fishery and Aquaculture Country Profiles Japan.* Rome: FAO. Retrieved 6 2018, from http://www.fao.org/fishery/facp/JPN/en.

[23] Fondriest Environment, Inc. (2016). Retrieved 6 2018, from Fundamentals of Environmental Measurements: http://www.fondriest.com/environmental–measurements/parameters/water–quality/turbidity–total–suspended–solids–water–clarity/.

[24] Forina, M., Casolino, C., & Almansa, E. M. (2003). The refinement of PLS models by iterative weighting of predictor variables and objects. *Chemometrics and Intelligent Laboratory Systems,* 68(1–2), 29–40.

[25] Forina, M., Lanteri, S., Cerrato Oliveros, M. C., & Pizarro Millan, C. (2004). Selection of useful predictors in multivariate calibration. *Analytical and Bioanalytical Chemistry,* 380(3), 397–418.

[26] Fukuyo, Y., Imai, I., Kodama, M., & Tamai, K. (2002). Red tides and other harmful blooms in Japan. In: Taylor, F.J.R., Trainer, V.L. (Eds.), *Harmful Algal Blooms in the PICES Region of the North Pacific. PICES Scientific Report No. 23* (pp. 7–20). Canada: North Pacific Marine Sci. Org. (PICES).

[27] Ghasemi, J., Niazi, A., & Leardi, R. (2003). Genetic–algorithm–based wavelength selection in multicomponent spectrophotometric determination by PLS: application on copper and zinc mixture. *Talanta,* 59(2), 311–317.

[28] Gholizadeh, M.H., Melesse, A.M., Reddi, L. (2016). A comprehensive review on water quality parameters estimation using remote sensing techniques. *Sensors*, 16, 1298.

[29] Giardino, C., Bresciani, M., Brag, F., Keukelaere, D. L., Knaeps, E., & Brando, V. (2017). Chapter 5 – Bio–optical Modeling of Total Suspended Solids. In D. R. Mishra, I. Ogashawara, & A. A. Gitelson (Eds.), *Bio–optical Modeling and Remote Sensing of Inland Waters* (pp. 129–156). Elsevier.

[30] Gitelson, A. A. (1992). The peak near 700 nm on radiance spectra of algae and water: relationships of its magnitude and position with chlorophyll concentration. *International Journal of Remote Sensing,* 13(17), 3367–3373.

[31] Gitelson, A. A. (1993). The nature of the peak near 700 nm on the radiance spectra and its application for remote estimation of phytoplankton pigments in inland waters, *Optical Engineering and Remote Sensing*, 170–179.

[32] Gitelson, A. A., Dall'Olmo, G., Moses, W., Rundquist, D. C., Barrow, T., Fisher, T. R., . . . Holz, J. (2008). Simple semi–analytical model for remote estimation of chlorophyll–a in turbid waters: Validation. *Remote Sensing of Environment,* 112(9), 3582–3593.

[33] Gitelson, A. A., Garbuzov, G., Szilagyi, F., Mittenzwey, K., Karnieli, A., & Kaiser, A. (1993). Quantitative remote sensing methods for real–time monitoring of inland waters quality. *Int. J. Remote Sens,* 14, 1269–1295.

[34] Gitelson, A. A., Gritz, Y., & Merzlyak, M. N. (2003). Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non–destructive chlorophyll assessment in higher plant leaves. *Journal of Plant Physiology,* 160(3), 271–282.

[35] Gitelson, A. A., Schalles, J. F., & Hladik, C. M. (2007). Remote chlorophyll–a retrieval in turbid, productive estuaries. *Remote Sens. Environ.,* 109, 464–472.

[36] Gons, H. J. (1999). Optical Teledetection of Chlorophyll a in Turbid Inland Waters. *Environ. Sci. Technol.,* 33(7), 1127–1132.

[37] Gordon, H. R. (1991). Absorption and scattering estimates from irradiance measurements: Monte Carlo simulations. *Limnology and Oceanography,* 36(4), pp. 769–777.

[38] Gordon, H. R., Brown, O. B., & Jacobs, M. M. (1975). Computed Relationships Between the Inherent and Apparent Optical Properties of a Flat Homogeneous Ocean. *Applied Optics,* 14(2), pp. 417–427.

[39] Gordon, H. R., Brown, O. B., Evans, R. H., Brown, J. W., Smith, R. C., & Baker, K. S. (1988). A semianalytic radiance model of ocean color. *Journal of Geophysical Research: Atmospheres,* 93(D9), pp. 10909–10924.

[40] Han, L., & Rundquist, D. C. (1997). Comparison of NIR/RED ratio and first derivative of reflectance in estimating algal–chlorophyll concentration: A case study in a turbid reservoir. *Remote Sensing of Environment,* 62(3), 253–261.

[41] Hashimoto, H., Hashimoto, T., Matsuda, O., Tada, K., Tamai, K., Uye, S., Yamamoto, T. (1997). Biological productivity of lower trophic levels of the Seto Inland Sea. In *Sustainable Development in the Seto Inland Sea, Japan – From the View Point of Fisheries*., Okaichi, T., Yanagi, T., Eds., Terra Scientific Publishing Company: Tokyo, Japan, pp. 17–58.

[42] Hu, C., Lee, Z., & Franz, B. (2012). Chlorophyll a algorithms for oligotrophic oceans: A novel approach based on three–band reflectance difference. *Journal of Geophysical Research: Oceans,* 117(C1).

[43] Hu, Z., Liu, H., Zhu, L., & Lin, F. (2011). Quantitative Inversion Model of Water Chlorophyll–a Based on Spectral Analysis. *Procedia Environmental Sciences,* 10, 523–528.

[44] Huang, Y., Jiang, D., Zhuang, D., & Fu, J. (2010). Evaluation of Hyperspectral Indices for Chlorophyll–a Concentration Estimation in Tangxun Lake (Wuhan, China). *Int J Environ Res Public Health,* 7(6), 2437–2451.

[45] Imai, I., Yamaguchi, M., Hori, Y. (2006). Eutrophication and occurrences of harmful algal blooms in the Seto Inland Sea, Japan. *Plankton Benthos Res,* 1 (2), 71–84.

[46] Inoue, Y., Peñuelas, J., Miyata, A., & Mano, M. (2008). Normalized difference spectral indices for estimating photosynthetic efficiency and capacity at a canopy scale derived from hyperspectral and CO2 flux. *Remote Sensing of Environment,* 112(1), 156–172.

[47] Inoue, Y., Sakaiya, E., Zhu, Y., & Takahashi, W. (2012). Diagnostic mapping of canopy nitrogen content in rice based on hyperspectral measurements. *Remote Sensing of Environment,* 126, 210–221.

[48] Irizuki, T., Hirose, K., Ueda, Y., Fujihara, Y., Ishiga, H., & Seto, K. (2018). Ecological shifts due to anthropogenic activities in the coastal seas of the Seto Inland Sea, Japan, since the 20th century. *Marine Pollution Bulletin,* 127, pp. 637–653.

[49] Katlane, R., Dupouy, C., Zargouni, F. (2012). Chlorophyll and turbidity concentration as an index of water quality of the Gulf of Gabes from MODIS in 2009. *Teledetection*, 11 (1), 263–271.

[50] Kawamura, K., Watanabe, N., Sakanoue, S., & Inoue, Y. (2008). Estimating forage biomass and quality in a mixed sown pasture based on partial least squares regression with waveband selection. *Grassland Science,* 54(3), 131–145.

[51] Kawamura, K., Watanabe, N., Sakanoue, S., Inoue, Y., & Odagawa, S. (2010). Testing genetic algorithm as a tool to select relevant wavebands from field hyperspectral data for estimating pasture mass and quality in a mixed sown pasture using partial least squares regression. *Grassland Science,* 56(4), 205–216.

[52] Khorram, S. (1982). Remote sensing of salinity in the San Francisco Bay Delta. *Remote Sensing*

of Environment, 12(1), 15–22.

[53] Kim, H., Son, S., Kim, Y. H., Khim, J. S., Nam, J., Chang, W. K., . . . Ryu, J. (2017). Remote sensing and water quality indicators in the Korean West coast: Spatio–temporal structures of MODIS–derived chlorophyll–a and total suspended solids. *Marine Pollution Bulletin, 121*(1–2), pp. 425–434.

[54] Kimura, H., Sato, M., Sugiyama, C., & Naganuma, T. (2001). Coupling of thraustochytrids and POM, and of bacterio– and phytoplankton in a semi–enclosed coastal area: implication for different substrate preference by the planktonic decomposers. *Aquatic Microbial Ecology*, 25(3), 293–300.

[55] Kirk, J. (1994). Light and photosynthesis in aquatic ecosystems, 2nd edition. *Cambridge University Press*, p. 509.

[56] Klemas, V. (2011). Remote Sensing of Sea Surface Salinity: An Overview with Case Studies. *Journal of Coastal Research,* 27(5), 830–838.

[57] Komorita, T., Guo, X., Yoshie, N., Fujii, N., & Takeoka, H. (2016). Temporal variation in water intrusion of a tidal frontal system and distribution of chlorophyll in the Seto Inland Sea, Japan. *Continental Shelf Research*, 112(1), 68–77.

[58] Kowalczuk, P., Stedmon, C. A., and Markager, S. (2006). Modelling abso rption by CDOM in the Baltic Sea from season, salinity and chlorophyll. *Mar. Chem.*, 101, 1–11.

[59] Kutser, T., Vahtmäe, E., Paavel, B., & Kauer, T. (2013). Removing glint effects from field radiometry data measured in optically complex coastal and inland waters. *Remote Sensing of Environment*, 133, 85–89.

[60] Lahet, F., Ouillon, S., & Forget, P. (2000). A Three–Component Model of Ocean Color and Its Application in the Ebro River Mouth Area. Remote Sensing of Environment, 72(2), 181–190.

[61] Latif, Z., Tasneem, M., Javed, T., Butt, S., Fazil, M., Ali, M., & Sajjad , M. (2003). Evaluation of Water–Quality by Chlorophyll and Dissolved Oxygen. *Water Resour. South Present Scenar. Future Prospect,* 7, 123–135.

[62] Lee, Z. P., Carder, K. L., Mobley, C. D., Steward, R. G., & Patch, J. S. (1998). Hyperspectral remote sensing for shallow waters. I. A Semianalytical model. *Applied Optics*, 37/27, 6329–6338.

[63] Lee, Z. P., Carder, K. L., Peacock, T. G., Davis, C. O., & Mueller, J. L. (1996). Method to derive ocean absorption coefficients from remote–sensing reflectance. *Applied Optics*, 35(3), pp. 453–462.

[64] Li, X., & He, Y. (2008). Chlorophyll Assessment and Sensitive Wavelength Exploration for Tea (Camellia sinensis) Based on Reflectance Spectral Characteristics. *HortScience,* 43(5), 1586–1591.

[65] Lu, F., Chen, Z., Liu, W., & Shao, H. (2016). Modeling chlorophyll–a concentrations using an artificial neural network for precisely eco–restoring lake basin. *Ecological Engineering,* 95, 422–

429.

[66] Marghany, M. (2010). Examining the Least Square Method to Retrieve Sea Surface Salinity from MODIS Satellite Data. *Eur. J. of Sci. Res.*, 40(3), 377–386.

[67] Ma, R., Ma, X., & Dai, J. (2007). Hyperspectral feature analysis of chlorophyll a and suspended solids using field measurements from Taihu Lake, eastern China. *Hydrol. Sci. J.,* 52(4), 808–824.

[68] Mateo–Sagasta , J., & Burke, J. (2010). *Agriculture and water quality interactions: a global overview.* Rome: FAO.

[69] McKee, D., Cunningham, A., & Jones, K. (1999). Simultaneous Measurements of Fluorescence and Beam Attenuation: Instrument Characterization and Interpretation of Signals from Stratified Coastal Waters. Estuarine, *Coastal and Shelf Science*, 48(1), 51–58.

[70] Mishra, S., & Mishra, D. (2012). Normalized difference chlorophyll index: A novel model for remote estimation of chlorophyll-a concentration in turbid productive waters. *Remote Sensing of Environment,* 117, 394–406.

[71] Mittenzwey, K. H., Breitwieser, S., Penig, J., Gitelson, A. A., Dubovitzkii, G., Garbusov, G., . . . Müller, A. (1991). Fluorescence and Reflectance for the *in situ* Determination of Some Quality Parameters of Surface Waters. *Acta hydrochimica et hydrobiologica,* 19(1), 1–15.

[72] Mittenzwey, K. H., Ullrich, S., Gitelson, A. A., Kondratiev, K. Y. (1992). Determination of chlorophyll a of inland waters on the basis of spectral reflectance. *Limnology and Oceanography*, 37, 147−149.

[73] Morel, A., & Prieur, L. (1977). Analysis of variations in ocean color. *Limnology and Oceanography,* 22(4), pp. 709–722.

[74] Moses, W.J., Gitelson, A.A., Perk, R.L., Gurlin, D., Rundquist, D.C., Leavitt, B.C., et al. (2012). Estimation of chlorophyll–a concentration in turbid productive waters using airborne hyperspectral data. *Water Research*, 46(4), 993–1004.

[75] NASA. (n.d.). Retrieved 6 2018, from OceanColor Web: https://oceancolor.gsfc.nasa.gov/atbd/chlor_a/

[76] Nechad, B., Ruddick, K. G., & Park, Y. (2010). Calibration and validation of a generic multisensor algorithm for mapping of total suspended matter in turbid waters. *Remote Sensing of Environment,* 114(4), 854–866.

[77] Nishijima, W., Umehara, A., Sekito, S., Okuda, T., Nakai, S. (2016). Spatial and temporal distributions of Secchi depths and chlorophyll a concentrations in the Suo Nada of the Seto Inland Sea, Japan, exposed to anthropogenic nutrient loading. *Science of the Total Environment*, 571, 543–550.

[78] Noh, J.H., Kim, W., Son, S.H., Ahn, J., Park, Y. (2018). Remote quantification of Cochlodinium polykrikoides blooms occurring in the East Sea using geostationary ocean color imager (GOCI). *Harmful Algae*, 73, 129–137.

[79] O'Reilly, J.E., and 24 Coauthors. (2000). SeaWiFS Postlaunch Calibration and Validation Analyses, Part 3. *NASA Tech. Memo. 2000–206892*, Vol. 11, S.B. Hooker and E.R. Firestone, Eds., NASA Goddard Space Flight Center, pp. 49.

[80] O'Reilly, J. E., Maritorena, S., Mitchell, B. G., Siegel, D. A., Carder, K. L., Garver, S. A., . . . McClain, C. (1998). Ocean color chlorophyll algorithms for SeaWiFS. *Journal of Geophysical Research: Oceans,* 103(C11), pp. 24937–24953.

[81] Oyama, Y., Matsushita, B., Fukushima, T., Matsushige, K., Imai, A. (2009). "Application of spectral decomposition algorithm for mapping water quality in a turbid lake (Lake Kasumigaura Japan) from Landsat TM data", *ISPRS Journal of Photogrammetry and Remote Sensing*, 64, 73−85.

[82] Pawar , V., Matsuda, O., Yamamoto , T., Hashimoto, T., & Rajendran, N. (2001). Spatial and temporal variations of sediment quality in and around fish cage farms: a case study of aquaculture in the Seto Inland Sea. *Fisheries Science,* 67(4), pp. 619–627.

[83] Pawar, V., Matsuda, O., & Fujisaki, N. (2002). Relationship between feed input and sediment quality of the fish cage farms. *Fisheries science,* 68(4), pp. 894–903.

[84] Pe´ rez, G., Queimalin˜os, C., Balseiro, E., Modenutti, B. (2007). Phytoplankton absorption spectra along the water column in deep North Patagonian Andean lakes (Argentina). *Limnologica*, 37, 3–16.

[85] Pope, R. M., & Fry, E. S. (1997). Absorption spectrum (380–700 nm) of pure water. II. Integrating cavity measurements. *Applied Optics,* 36(33), pp. 8710–8723.

[86] Pulliainen, J., Kallio, K., Eloheimo, K., Koponen, S., Servomaa, H., Hannonen, T., . . . Hallikainen, M. (2001). A semi–operative approach to lake water quality retrieval from remote sensing data. *Science of The Total Environment,* 268(1–3), 79–93.

[87] Qing, S., Zhang, J., Cui, T., & Bao, Y. (2013). Retrieval of sea surface salinity with MERIS and MODIS data in the Bohai Sea. *Remote Sensing of Environment,* 136, pp. 117–125.

[88] Rönnberg, C., & Bonsdorff, E. (2004). Baltic Sea eutrophication: area–specific ecological consequences. *Hydrobiologia,* 514(1–3), 227–241.

[89] Ryan, K., & Ali, K. (2016). Application of a partial least–squares regression model to retrieve chlorophyll–a concentrations in coastal waters using hyper–spectral data. *Ocean Science Journal,* 51(2), 209–221.

[90] Sakuno, Y., Makio, K., Koike, K., Thaw, M.–S.–H., & Kitahara, S. (2013). Chlorophyll–a Estimation in Tachibana Bay by Data Fusion of GOCI and MODIS Using Linear Combination Index Algorithm. *Advances in Remote Sensing,* 2(4), 292–296.

[91] Sakuno, Y., Miño, E. R., Nakai, S., Mutsuda, H., Okuda, T., Nishijima, W., Castro, R., García, A., Peña, R., Rodríguez, M., Depratt, G. C. (2014). Chlorophyll and suspended sediment mapping to the Caribbean Sea from rivers in the capital city of the Dominican Republic using ALOS AVNIR–

2 data. *Environ. Monit. Assess.*, 186(7), 4181–4193.

[92] Sasaki, H., Tanaka, A., Iwataki, M., Touke, Y., Siswanto, E., Tan, C. K., Ishizaka, J. (2008). Optical properties of the red tide in Isahaya Bay, south–western Japan: influence of chlorophyll a concentration. *J. Oceanogr.*, 64, 511–523.

[93] Shafique, N., Fulk, F., Autrey, B., & Flotemersch, J. (2003). Hyperspectral Remote Sensing of Water Quality Parameters for Large Rivers in the Ohio River Basin. *In Proceedings of the First Interagency Conference on Research in the Watersheds* (pp. 27–30). Washington, DC: USDA Agricultural Research Service.

[94] Shanmugam, P., He, X., Singh, R. K., & Varunan, T. (2018). A modern robust approach to remotely estimate chlorophyll in coastal and inland zones. *Advances in Space Research,* 61(10), pp. 2491–2509.

[95] Siddorn, J. R., Bowers, D. G., & Hoguane, A. M. (2001). Detecting the Zambezi River Plume using Observed Optical Properties. *Marine Pollution Bulletin*, 42(10), 942–950.

[96] Sikorska, A., Giudice, D., Banasik, K., & Rieckermann, J. (2015). The value of streamflow data in improving TSS predictions – Bayesian multi–objective calibration. *Journal of Hydrology,* 530, 241–254.

[97] Smith, R. C., & Baker, K. S. (1981). Optical properties of the clearest natural waters (200–800 nm). *Applied Optics*, 20(2), 177–184.

[98] Song, K., Li, L., Li, S., Tedesco, L., Duan, H., Li, Z., . . . Shao, T. (2014). Using Partial Least Squares–Artificial Neural Network for Inversion of Inland Water Chlorophyll–a. *IEEE Transactions on Geoscience and Remote Sensing,* 52(2), 1502–1517.

[99] Song, K., Li, L., Tedesco, L. P., Li, S., Clercin, N., Li, Z. C., et al. (2012). Hyperspectral determination of eutrophication for a water supply source via genetic algorithm–partial least square (GA–PLS) modeling. *Science of the Total Environment*, 426, 220–232.

[100] Song, K., Li, L., Tedesco, L. P., Li, S., Duan, H., Liu, D., . . . Zhao, Y. (2013). Remote estimation of chlorophyll–a in turbid inland waters: Three–band model versus GA–PLS model. *Remote Sensing of Environment,* 136, 342–357.

[101] Stagakis, S., Markos, N., Sykioti, O., & Kyparissis, A. (2010). Monitoring canopy biophysical and biochemical parameters in ecosystem scale using satellite hyperspectral imagery: An application on a Phlomis fruticosa Mediterranean ecosystem using multiangular CHRIS/PROBA observations. *Remote Sensing of Environment,* 114(5), 977–994.

[102] Stratoulias, D., Balzter, H., Zlinszky, A., & Tóth, V. R. (2015). Assessment of ecophysiology of lake shore reed vegetation based on chlorophyll fluorescence, field spectroscopy and hyperspectral airborne imagery. *Remote Sensing of Environment,* 157, 72–84.

[103] Stumpf, R. P. (2001). Applications of Satellite Ocean Color Sensors for Monitoring and Predicting Harmful Algal Blooms. *Journal of Human and Ecological Risk Assessment*, 7,

1363−1368.

[104] Sugihara, S., Kishino, M., & Okami, N. (1985). Estimation of water quality parameters from irradiance reflectance using optical models. *Journal of the Oceanographical Society of Japan,* 41(6), 399–406.

[105] Swierenga, H., De Groot, P., De Weijer, A., Derksen, M., & Buydens, L. (1998). Improvement of PLS model transferability by robust wavelength selection. *Chemometrics and Intelligent Laboratory Systems,* 41(2), 237–248.

[106] Tomlinson, M. C., Stumpf, R. P., Ransibrahmanakul, V., Truby, E. W., Kirkpatrick, G. J., Pederson, B. A., et al. (2004). Evaluation of the use of SeaWiFS imagery for detecting Karenia brevis harmful algal blooms in the eastern Gulf of Mexico. *Remote Sensing of Environment*, 91, 293−303.

[107] UNESCO. (2015). *International Initiative on Water Quality.* Paris: UNESCO.

[108] Voutilainen, A., Pyhälahti, T., Kallio, K., Pulliainen, J., Haario, H., & Kaipio, J. (2007). A filtering approach for estimating lake water quality from remote sensing data. *International Journal of Applied Earth Observation and Geoinformation,* 9(1), 50–64.

[109] Wang, X. J., Liu, R. M. (2005). Spatial analysis and eutrophication assessment for chlorophyll a in Taihu Lake. *Environmental Monitoring and Assessment*, 101(1/2/3), 1–8.

[110] Wang, Z., Kawamura, K., Sakuno, Y., Fan, X., Gong, Z., Lim, J. (2017). Retrieval of chlorophyll–a and total suspended solids using iterative stepwise elimination partial least squares (ISE–PLS) regression based on field hyperspectral measurements in irrigation ponds in Higashihiroshima, Japan. *Remote Sens.*, 9, 264.

[111] WHO. (2011). *Guidelines for drinking–water quality, 4th edition.* Geneva: WHO.

[112] Williams, P. C. (2001). Implementation of Near–Infrared Technology. In P. Williams, & K. Norris (Eds.), *Near–Infrared Technology in the Agricultural and Food Industries, 2nd ed.* (pp. 145–170). St. Paul, Minnesota, USA: American Association of Cereal Chemists Inc.

[113] Wold, H. (1966). Estimation of Principal Components and Related Models by Iterative Least squares. In P. Krishnaiaah (Ed.), *Multivariate Analysis* (pp. 391–420). New York: Academic Press.

[114] Wong, M. S., Lee, K. H., Kim, Y. J., Nichol, J. E., Li, Z., & Emerson, N. (2007). Modeling of Suspended Solids and Sea Surface Salinity in Hong Kong using Aqua/MODIS Satellite Images. *Korean Journal of Remote Sensing,* 23(3), 161–169.

[115] Yacobi, Y. Z., Moses, W. J., Kaganovsky, S., Sulimani, B., Leavitt, B. C., & Gitelson, A. A. (2011). NIR–red reflectance–based algorithms for chlorophyll–a estimation in mesotrophic inland and coastal waters: Lake Kinneret case study. *Send to,* 45(7), 2428–2436.

[116] Yamamoto, T. (2003). The Seto Inland Sea – eutrophic or oligotrophic, Mar. Poll. Bull., 47, 37–42.

[117] Yang, X., Wu, X., Hao, H., & He, Z. (2008). Mechanisms and assessment of water eutrophication.

J Zhejiang Univ Sci B., 9(3), 197–209.Pope

[118] Yoshie, N., Guo, X., Fujii, N., Komorita, T. Ecosystem and nutrient dynamics in the Seto Inland Sea, Japan. (2011). In *Interdisciplinary Studies on Environmental Chemistry–Marine Environmental Modelling and Analysis*,, Omori, K., Guo, X., Yoshie, N., Fujii, N., Handoh, I.C., Isobe, A., Tanabe, S. Eds., Terrapub, Tokyo, Japan, pp. 39–49.

[119] Yunus, A. P., Dou, J., Sravanthi, N. (2015). Remote sensing of chlorophyll–a as a measure of red tide in Tokyo Bay using hotspot analysis. *Remote Sens. Appl. Soc. Environ.*, 2, 11–25.

[120] Zheng, G., & DiGiacomo, P. (2017). Remote sensing of chlorophyll–a in coastal waters based on the light absorption coefficient of phytoplankton. *Remote Sensing of Environment,* 201, pp. 331–341.

[121] Zhao, J., Ghedira, H. (2014). Monitoring red tide with satellite imagery and numerical models: a case study in the Arabian Gulf. *Mar. Pollut. Bull.*, 79, 305–313.

Appendix A. The summary of collected water quality data in irrigation ponds in Higashihiroshima, Japan.

Date	Station	Chl-*a* (μg/L)	TSS (mg/L)
29 Sep, 2013	Stn1	78.6428	10.4
	Stn2	56.1675	33.2
	Stn3	3.3690	0.5
	Stn4	3.1445	0.6
	Stn5	6.3463	2.6
	Stn6	2.1693	0.7
31 Oct, 2013	Stn1	44.2312	40.7
	Stn2	51.6339	23.4
	Stn3	2.2973	1.5
	Stn4	2.9538	0.9
	Stn5	1.7559	4.5
	Stn6	2.5810	1.8
16 Nov, 2013	Stn1	5.6515	8.7
	Stn2	54.3698	31.7
	Stn3	1.6020	0.8
	Stn4	1.7560	0.3
	Stn5	2.4143	2.9
	Stn6	0.8688	1.6
2 Dec, 2013	Stn1	6.5835	2.8
	Stn2	42.8922	20.9
	Stn3	1.6245	0.1
	Stn4	0.5906	0.6
	Stn5	1.3536	3.3
	Stn6	0.9725	1.8
3 Jan, 2014	Stn1	98.7437	16.8
	Stn2	21.6265	13.7
	Stn3	0.5053	0.1
	Stn4	1.2877	0.9
	Stn5	1.7007	3.6
	Stn6	0.0862	1.5
19 Jan, 2014	Stn1	169.5454	26.5
	Stn2	42.7523	15.6
	Stn3	0.8687	0.1
	Stn4	1.6688	0.1
	Stn5	0.9734	2.2
	Stn6	0.1048	0.8
24 Mar, 2014	Stn1	169.0605	19.2
	Stn2	46.6705	38
	Stn3	2.6385	0.9
	Stn4	0.5793	0.4
	Stn5	1.9775	1.7
	Stn6	0.0000	0.9
9 Apr, 2014	Stn1	0.5475	1.2
	Stn2	48.5398	33.50
	Stn3	0.6453	0.9
	Stn4	0.6748	0.5
	Stn5	1.2925	1.4
	Stn6	0.6712	1.4
24 May, 2014	Stn1	11.9524	5.2
	Stn2	37.7392	26
	Stn3	0.8618	1.1
	Stn4	1.0956	1.1
	Stn5	1.6796	1.3
	Stn6	1.7512	0.2
28 Jun, 2014	Stn1	13.4963	4.80
	Stn2	133.9240	53.00
	Stn3	2.7226	0.80
	Stn4	7.9958	2.50
	Stn5	3.1450	1.20
	Stn6	1.6008	0.30

Appendix A. The summary of collected water quality data in irrigation ponds in Higashihiroshima, Japan.

Date	**Station**	**Chl-*a* (μg/L)**	**TSS (mg/L)**
15 Jul, 2014	Stn1	16.1804	11.80
	Stn2	157.7630	83.6
	Stn3	2.4173	0.70
	Stn4	5.4253	1.80
	Stn5	1.4901	1.30
	Stn6	1.9517	2.00

Station	**Pond**	**Coordinate X**	**Coordinate Y**
Stn1	Nanatsu-ike	132°41'39" E	34°26'06"N
Stn2	Shitami-Oike	132°42'22"E	34°24'28"N
Stn3	Okuda-Oike	132°43'43"E	34°24'25"N
Stn4	Yamanaka-ike	132°43'12"E	34°24'14"N
Stn5	Yamanakaike-kamiike	132°43'14"E	34°24'15"N
Stn6	Budou-ike	132°42'45"E	34°24'02"N

Appendix B. The summary of collected water quality data in the Seto Inland Sea, Japan.

Date	Station	Chl-*a* (μg/L)	SSS (psu)	SST (℃)
16 Dec, 2015	Stn1	3.95	31.26	15.4
	Stn2	5.27	31.28	14.89
	Stn3	5.7	31.48	15.05
	Stn4	13.33	30.75	14.61
	Stn5	3.78	30.51	14.55
	Stn1-2	5.65	31.25	14.93
	Ushinokubi 1	4.16	31.39	14.99
13 Jan, 2016	Stn1	3.15	32.47	12.97
	Stn2	6.72	32.04	11.94
	Stn3	5.11	31.94	11.72
	Stn4	4.94	30.99	10.95
	Stn5	4.12	32.04	11.74
	Stn6	3.83	31.92	11.83
	Stn1-2	3.83	11.47	12.52
	Ushinokubi 1	7.06	31.91	11.68
	Ushinokubi 2	5.86	31.9	11.69
22 Mar, 2016	Stn1	1.84	32.8	11.69
	Stn2	1.39	30.47	11.54
	Stn3	2.46	31.94	11.04
	Stn4	3.5	31.71	11.24
	Stn5	1.75	32.27	11.26
	Stn6	2.99	31.99	11.42
	Ushinokubi 1	1.04	32.47	12.16
	Ushinokubi 2	2.86	32.45	11.66
14 Jun, 2016	Stn1	2.38	32.12	20.67
	Stn2	5.49	30.55	21.37
	Stn3	6.11	29.48	22.23
	Stn4	14.33	21.99	23.27
	Stn5	3.73	30.22	22.76
	Stn6	5.3	31.08	21.61

Appendix B. The summary of collected water quality data in the Seto Inland Sea, Japan.

Date	Station	Chl-*a* (μg/L)	SSS (psu)	SST (℃)
13 Oct, 2016	Stn1	6.77	30.78	24.65
	Stn2	11.08	30.55	24.24
	Stn3	17.6	30.09	21.61
	Stn4	16.42	29.78	23.72
	Stn5	23.34	30.13	23.97
	Stn6	12.6	30.26	24.11
21 Dec, 2016	Stn1	3.24	31.86	15.54
	Stn2	6.11	31.46	14.44
	Stn3	7.84	30.94	13.65
	Stn4	13.94	30.34	13.2
18 Jan, 2017	Stn1	3.05	32.35	12.59
	Stn2	6.48	32.12	11.2
	Stn3	5.15	31.98	10.74
	Stn4	6.22	31.69	10.58
	Stn5	4.69	32.33	11.61
	Stn6	8.74	32.19	11.24
	Stn1-2	12.15	32.31	11.89
	Stn7	2.76	32.11	11.07
15 Mar, 2017	Stn1	0.83	33.11	11.1
	Stn2	1.06	32.85	10.72
	Stn3	1.72	32.65	10.43
	Stn4	3.9	31.18	10.27
	Stn5	5.46	32.41	10.22
	Stn6	1.2	32.73	10.24
	Stn1-2	0.73	32.94	10.88
	Stn7	0.53	32.69	10.61
18 Apr, 2017	Stn1	2.86	32.67	14.1
	Stn2	4.22	32.54	14.08
	Stn3	2.67	32.37	14.01
	Stn4	5.01	30.98	15.09
	Stn5b	5.07	25.38	14.77

Date	**Station**	**Chl-*a* (µg/L)**	**SSS (psu)**	**SST (℃)**
16 May, 2017	Stn1	2.83	33.16	16.72
	Stn2	2.43	32.99	17.11
	Stn3	4.66	32.51	17.22
	Stn4	7.68	32.2	18.15
	Stn5b	5.85	32.37	17.75
18 Jul, 2017	Stn1	2.82	32.66	23.75
	Stn2	2.22	32.59	24.71
	Stn3	3.65	32.54	24.01
	Stn4	2.31	32.35	24.19

Date	**Station**	**Chl-*a* (µg/L)**	**SSS (psu)**	**SST (℃)**
16 Aug, 2017	Stn1	1.64	32.89	27.48
	Stn2	1.26	32.61	27.57
	Stn3	4.06	32.41	27.14
	Stn4	8.82	31.29	28.14
	Stn5b	2.16	32.47	27.14

Station	**Coordinate X**	**Coordinate Y**
Stn1	133°15' 24" E	34°19' 44" N
Stn2	133°19' 26" E	34°20' 31" N
Stn3	133°22' 10" E	34°21' 51" N
Stn4	133°24' 44" E	34°24' 37" N
Stn5	133°24' 58" E	34°22' 01" N
Stn6	133°27' 58" E	34°23' 38" N

Station	**Coordinate X**	**Coordinate Y**
Ushinokubi 1	133°19' 30" E	34°21' 56" N
Ushinokubi 2	133°19' 38" E	34°21' 45" N
Stn1-2	133°18' 6" E	34°20' 40" N
Stn5b	133°16' 11" E	34°20' 11" N
Stn7	133°20' 24" E	34°22' 8" N