Development of A Semi-Analytical Algorithm for Estimating Water Transparency in Various Waters from MERIS Data

July 2022

Msusa Anastazia Daniel

Development of A Semi-Analytical Algorithm for Estimating Water Transparency in Various Waters from MERIS Data

A Dissertation Submitted to
the Graduate School of Life and Environmental Sciences,
the University of Tsukuba
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in Science
(Doctoral Program in Integrative Environment and Biomass Sciences)

Msusa Anastazia Daniel

Abstract

Water transparency (Secchi disk depth; Z_{SD}) is a key parameter of water quality and thus it is very important to routinely monitor. Recently, remote sensing technique has been widely used because of its capability of giving synoptic view of Z_{SD} distribution in waters as well as the progress of estimation algorithm development. Previous studies have reported that the Z_{SD} estimation algorithms based on the new theory perform better than those algorithms based on the classic theory. In the new theory-based algorithms, accurate estimation of absorption and backscattering coefficients (a and b_b) using an appropriate quasi-analytical algorithm (QAA) is a key step for estimating more accurate Z_{SD} values. Therefore, the selection of appropriate QAA for different water types is necessary.

Consequently, the objectives of this study were to: (1) develop a new algorithm for estimating more accurate Z_{SD} by selecting more appropriate QAA for different water types; (2) evaluate the performance of the developed Z_{SD} estimation algorithm using synthetic and *in situ* Z_{SD} values; and (3) apply the developed Z_{SD} estimation algorithm to Medium Resolution Imaging Spectrometer (MERIS) data over Lake Kasumigaura between 2003 and 2012 to produce a long-term Z_{SD} product.

In Chapter 2, I made four efforts to improve a state-of-art Z_{SD} estimation algorithm, which is based on the new underwater visibility theory. First, I adopted a recently proposed water type classification algorithm, which can classify all waters into four types, i.e., clear (Type I), moderately turbid (Type II), highly turbid (Type III), and extremely turbid (Type IV). This effort can promise an assumption for absorption coefficients at reference bands

become more valid. Second, I selected three semi-analytical models and one empirical model specific to each water type to retrieve accurate a and b_{b} at reference bands. Third, I found two empirical models and one semi-analytical model, which can represent reasonable shapes for particulate backscattering coefficients, and thus to retrieve accurate a and b_{b} at a given wavelength for each water type. Fourth, I carried out a statistical analysis using a synthetic dataset to identify possibility of minimum K_{d} for each water type. This effort can avoid selection of inappropriate wavelengths due to estimation errors in the previous steps. Validation results obtained from the synthetic dataset (N=91287, Z_{SD} values ranging from 0.01m to 44.68m) show that the developed Z_{SD} algorithm in this study outperformed the state-of-art algorithm with a Root Mean Square Error (RMSE) reduced from 0.30 to 0.23, Mean Absolute Percentage Error (MAPE) reduced from 116% to 65%. Particularly, the new algorithm does not include estimation outliers, which were obviously observed in Z_{SD} estimations by using the previous algorithm.

In Chapter 3, I carried out two more validations for the developed Z_{SD} estimation algorithm by using an independently generated synthetic dataset (N=1000, Z_{SD} ranging from 0.01m to 34.87m) and *in situ*-collected Z_{SD} measurements (N=305, Z_{SD} ranging from 0.3m to 16.4m). The performance of the developed Z_{SD} algorithm was also compared to that of the state-of-art algorithm. Results from the synthetic dataset show a significant improvement with RMSE of 0.21 versus 0.16 and MAPE of 52% versus 38%. However, a systematic overestimation was still observed (bias=35%). This systematic overestimation is considered to be caused by the assumption for the absorption coefficients at the reference wavelengths. On the other hand, comparison results from *in situ* data shows the improvement in accuracy

with RMSE of 0.18 versus 0.17 and MAPE of 32% versus 27%. In addition, the systematic overestimation was not found for in situ data. This is probably because that the measurement errors contained in *in situ* data have hidden the systematic overestimation due to the mechanism of the developed algorithm. Overall, all results show that the four improvements described in Chapter 2 worked, and thus increased Z_{SD} estimation accuracy.

In Chapter 4, I applied the developed Z_{SD} algorithm to the 2003-2012 MERIS images collected from Lake Kasumigaura, Japan to check the performance of the algorithm. From 19 matchups, the results showed good performance with RMSE of 0.08, MAPE of 15%, and bias of -9%. The satellite retrieved long-term Z_{SD} trends agreed well with that retrieved from the *in situ* long-term Z_{SD}. These results indicate that satellite data can be used for Z_{SD} monitoring.

Acknowledgement

I would like to express my gratitude to my supervisor, Prof. Matsushita Bunkei, for all of his guidance, wisdom and full support throughout this study, from inception to completion, as well as his incredible and unwavering patience in working with me over the past four years. I really appreciate his constructive ideas and comments, availability to discuss about my research and well read my manuscript. His shared skills and knowledge in the field of remote sensing are very useful to my future work.

My sincere gratitude to the rest of my thesis committee, Prof. Fukushima Takehiko, Prof. Nasahara Kenlo and Prof. Kato Hiroaki, for their valuable comments, suggestions, and questions.

I am deeply thankful to Dr. Dalin Jiang for his constructive suggestions and providing Synthetic data that have been very helpful in this research. My sincere thanks to Mrs. Suzuki Mika for her help, encouragement and support in the laboratory and administrative work. I would like to thank all Laboratory members for their valuable contribution in my research especially during seminars.

I would like to thank those who provided data and make this research possible: the European Space Agency (ESA) for providing the MERIS satellite images, the National Institute for Environmental Studies, Japan (NIES) for providing the *in situ* data for Lake Kasumigaura.

Special thanks to the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan for providing the scholarship. I am highly indebted to the Study in Japan team and Embassy of Japan in Tanzania in general for facilitating the entire process of

acquiring the MEXT scholarship.

I would like to express my special gratitude towards my family for the support, encouragement and prayers that helped me in completion of this study. My beloved and supportive husband Octavian being always by my side, my beloved son Olney who served as my inspiration to pursue this undertaking.

Contents

List of Tables

List of Figures

Chapter I: General Introduction

1.1. Water transparency

Water transparency relates to the depth that light will penetrate water (Figure 1.1.), and thus it is key information for water quality evaluation (Swift et al., 2006; Doron et al., 2007; Fleming-Lehtinen et al., 2012; Lee et al., 2015; Fukushima et al., 2016; Rodrigues et al., 2017; Liu et al., 2020; Olmanson et al., 2008). Water transparency depends on the amount of particles (which can be algae or sediment from erosion) present in the water whereby the more the particles, the less the water transparency. Higher water transparency may indicate lower levels of turbidity and vice versa while lower water transparency can cause water temperature to rise.

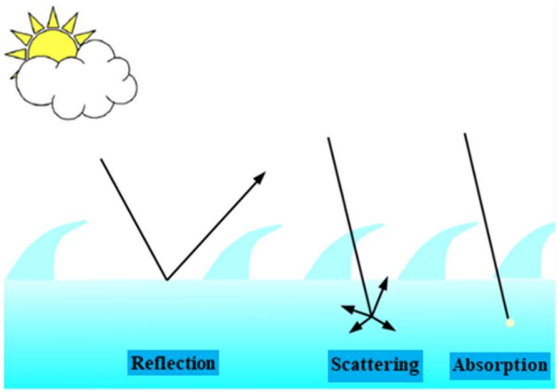


Figure 1.1. Light paths in water

Estimating water transparency is important because the clarity of water impacts the amount of light penetration which affects photosynthesis, on the other hand water transparency is used for comparing bodies of water or looking for changes to a specific body

of water over time (Wondie et al., 2007; Rodrigues et al., 2017). Changes in clarity of water can be an indicator of a human threat to an ecosystem (Canfield & Langeland, 1985; Lee et al., 1999; Holland, 1993; Devlin et al., 2008; McCullough et al, 2012).

1.2. Measurement of water transparency

In the field survey of aquatic environment water transparency has been a routine measurement since 1860s (Lee et al., 1995; Sandén & Håkansson, 1996; Fleming-lehtinen et al., 2016). Water transparency is often represented using Secchi disk depth (Z_{SD}), which can be measured using either traditional Secchi disk or remote sensing technique.

1.2.1. Tradition method

Tradition measurement of water transparency involves the use of a Secchi disk (which is a white or black-and-white disk) that is lowered into the water on a cord (Tyler, 1968). Z_{SD} measurement in the field is the depth at which the observer can no longer see the disk and where it reappears again when it is raised (Figure 1.2.).

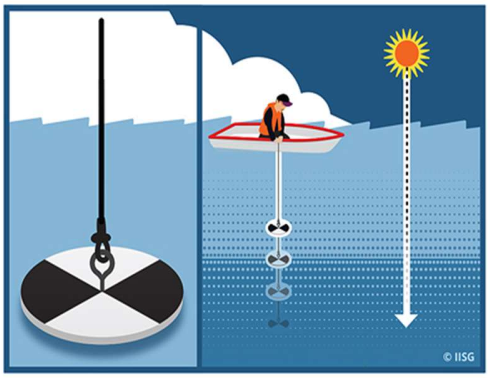


Figure 1.2. Z_{SD} Measurements in the field.

For decades now, Secchi depth measurements has been practiced across a wide geographic area because of its reliability, easy and affordability nature in water clarity measurement. The measurements of Secchi depth are typically done in deeper areas of the lake and at regular intervals ideally in multiple locations. These measurements can vary from a few centimeters in very turbid lakes to more than 10 meters in deep, clear lakes (Tyler, 1968; Sandén & Håkansson, 1996; Wernand, 2010).

However, collection of *in situ* Secchi depth over a wide area is characterized by temporal and spatial coverage limitation, laborious in nature, time-consuming and lacks data in some waters necessitating other methods to expand the temporal and spatial scope of water quality assessment (Lee et al., 2015; Mao et al., 2018; Wang et al., 2019).

1.2.2. Remote sensing techniques

The ultimate goal of using remote sensing technique is to avoid relying on *in situ* data collection as the only source of water quality data, and to base on remote sensing techniques because of its capability of giving a synoptic view of Z_{SD} distribution in waters as well as the progress of estimation algorithm development (Bai et al., 2020; Sòria-Perpinyà et al., 2020; Chang et al., 2020). There are also some limitations when using remote sensing techniques; for example, too large pixel size for small water bodies, too long temporal return period to effectively capture rapid changes in water surface conditions and poor atmospheric condition (Lee et al., 2015; Mao et al., 2018; Wang et al., 2019).

However, remote sensing techniques have been acknowledged as an effective tool for monitoring trends in Secchi depth, so data obtained from remote sensing have been used to

evaluate water transparency in lakes (Bai et al., 2020; Sòria-Perpinyà et al., 2020) . Therefore, transferability of models among locations and times is very important aspect to the application of remote sensing methods to lake monitoring (Sòria-Perpinyà et al., 2020; Chang et al., 2020). Generally, there are two types of Z_{SD} estimation algorithms: (1) empirical algorithm, and (2) semi-analytical algorithm.

1.2.2.1. Empirical Algorithms

Empirical algorithms use a simple regression analysis between remote sensing data and *in situ*-measured Z_{SD} values (Lee et al., 2018; Alikas & Kratzer, 2017; Canfield & Langeland, 1985; Doron et al., 2011). These algorithms are usually simple but require *in situ* data for recalibration before applying to another water body. For this reason, waters without *in situ* data, the empirical algorithms applicability is a challenge.

1.2.2.2. Semi-analytical Algorithms

Semi-analytical algorithms are based on an underwater visibility theory (Lee et al., 2015; Mao et al., 2018; Wang et al., 2019; Jiang et al., 2019). Despite of Semi-analytical algorithms being complicated and their assumptions depending on optical properties of a water body, they are more suitable for estimating Z_{SD} values from remote sensing data because they can be applied to various waters without the requirements of recalibration when compared to the empirical algorithms (Vundo et al., 2019; Jiang et al., 2019). For this reason, Semi-analytical algorithms became more useful for Z_{SD} monitoring especially in those waters which lack *in situ* data.

There are two underwater visibility theories until now for retrieving Z_{SD}: (1) classic theory proposed in 1952 and (2) new theory proposed in 2015. Classical theory are semi-analytical algorithms for Z_{SD} retrieval which are based mainly on an underwater visibility theory proposed by Duntley (1952). Among all semi-analytical algorithms of Z_{SD} retrieval basing on classical theory, the algorithm proposed by Doron et al. (2011) performed better than other algorithms (Lee et al., 2015). In classical theory, Z_{SD} is inversely proportional to the sum of K_d (diffuse attenuation coefficient of downwelling irradiance (m^{-1})) and c(beam attenuation coefficient (m^{-1})) within the visible domain (Tyler, 1968; Sandén & Håkansson, 1996; Wernand, 2010).

$$Z_{SD} = \frac{\Gamma}{K_d(v) + c(v)} \tag{1.1}$$

$$\Gamma = \ln\left(\frac{R_t - R_w}{c_t R_w}\right) \tag{1.2}$$

Whereby, $K_d(v)$ is the diffuse attenuation coefficient of downwelling irradiance (m^{-1}) in visible domain, $c(v)$ is the beam attenuation coefficient (m^{-1}) in visible domain and Γ is the coupling constant with a range of 5-10. Γ can be estimated using equation. (1.2), where, R_t is the submerged disk reflectance, R_w is the reflectance of water body around the disk, and c_t is the threshold of contrast for detecting a disk.

Lee et al. (2015) pointed out some remaining problems in the classic theory. First, there is no universal relationship between Z_{SD} and the beam attenuation coefficient ($c(v)$). Second, the correlation between Z_{SD} and the diffuse attenuation coefficient of downwelling irradiance

(K_d) is typically similar or better than the correlation between Z_{SD} and the beam attenuation coefficient (c(v)).

To tackle the above shortcomings, Lee et al. (2015) developed a new theory for underwater visibility. Previous studies have reported that the Z_{SD} estimation algorithms based on the new theory perform better than those algorithms based on the classic theory (Lee et al., 2015; Vundo et al., 2019; Jiang et al., 2019; Wang et al., 2019; Bowers et al., 2020). The new theory-based algorithms (hereafter referred to as the 'Lee15 algorithm') rely on diffuse attenuation coefficient (K_d) at a wavelength corresponding to the maximum transparency for such interpretation (Lee et al., 2015).

$$Z_{SD} = \frac{1}{2.5 Min K_d(\lambda)} \ln\left(\frac{\left|0.14 - R_{rs}^{PC}\right|}{C_t^r}\right) \quad (1.3)$$

Where $MinK_d(\lambda)$ is the minimum diffuse attenuation coefficient of downwelling irradiance ($K_d(\lambda)$) value in the visible domain. R_{rs}^{PC} is the corresponding remote sensing reflectance (Rrs, sr^{-1}) at the band with the minimum $K_d(\lambda)$, and C_t^r is the contrast threshold for sighting a white disk by eyes (which is equal to 0.013 sr^{-1}). The coefficient of 2.5 was obtained with the assumption that $K_T = 1.5K_d$, while K_T is the upwelling radiance diffuse attenuation coefficient (m^{-1}).

Recently, Jiang et al. (2019) listed some shortcomings in Lee15 algorithm towards more accurate Z_{SD} values retrieval. First, the estimation of absorption coefficient (a) and backscattering coefficient (b_b) using QAA_V6 or QAA_V5 (Quasi-Analytical Algorithm

Version 6 or 5) failed in turbid inland waters. Second, using a fixed value of K_T/K_d ratio in Z_{SD} values estimation may lead to some errors.

So, Jiang et al. (2019) developed a new theory-based semi-analytical algorithm (hereafter refered to as Jiang19 algorithm) to estimate Z_{SD} values with the aim of improving the former Z_{SD} estimation algorithm developed by Lee et al. (2015). The improvement was done based on using QAA_hybrid (which integrates different types of QAAs) and dynamic K_T/K_d ratio instead of fixed ratio of 1.5.

The Jiang19 algorithm contains three main steps. First, a QAA hybrid is developed to retrieve $a(\lambda)$ and $b_b(\lambda)$ from $R_{rs}(\lambda)$ spectra. In the QAA hybrid, a maximum chlorophyll index (MCI) proposed by Gower et al., (2005) is used to switch QAA_V5 (Lee et al., 2002) and QAA_T (i.e., QAA_Turbid in (Yang et al., 2013)). In detail, if MCI <= 0.0016 sr^{-1}, 560 nm is used as the reference wavelength and thus QAA_V5 is selected to estimate $a(\lambda)$ and $b_b(\lambda)$ for clear waters; otherwise, 754 nm is used as the reference wavelength and thus QAA_T is selected to estimate $a(\lambda)$ and $b_b(\lambda)$ for turbid waters.

Then, $K_d(\lambda)$ is estimated from $a(\lambda)$ and $b_b(\lambda)$ using Equation (1.4) developed by Lee et al. (2005, 2013) as the second step, and Z_{SD} is estimated from the minimum $K_d(\lambda)$ in the visible domain (e.g., 443 nm, 490 nm, 510 nm, 560 nm, 620 nm, 665 nm) and the corresponding $R_{rs}(\lambda)$ using equation (1.5) (the third step).

$$K_d(\lambda) = (1 + 0.005\theta)a(\lambda) + 4.259(1 - 0.265\eta_w(\lambda)\left(1 - 0.52e^{-10.8a(\lambda)}\right)b_b(\lambda) \qquad (1.4)$$

$$Z_{\mathrm{SD}} = \frac{1}{(1 + K_{\mathrm{T}}/K_{\mathrm{d}}).\,\mathrm{Min}K_{\mathrm{d}}(\lambda)} \ln\left(\frac{|0.14 - R_{\mathrm{rs}}^{\mathrm{PC}}|}{C_{\mathrm{t}}^{r}}\right), \quad K_{\mathrm{T}}/K_{\mathrm{d}} = \frac{1.04(1 + 5.4\mu)^{0.5}}{1/\left(1 - \frac{sin(\theta)^2}{RI^2}\right)^{0.5}} \qquad (1.5)$$

where θ is the solar zenith angle and η_{w} is the ratio of b_{bw} (backscattering coefficient of pure water (Zhang & Hu, 2009)) to b_{b}. RI is the refractive index value of water (1.34, (Lee et al., 1998)), K_{T} is the diffuse attenuation coefficient of upwelling radiance at the wavelength with the minimum $K_{\mathrm{d}}(\lambda)$, and μ is defined as $b_{\mathrm{b}}/(a + b_b)$.

1.3. Remaining Issues

In the new theory algorithms (both Lee15 and Jiang19 algorithms), accurate estimation of absorption and backscattering coefficients (a and b_b) using an appropriate quasi-analytical algorithm (QAA) is a key step for estimating more accurate Z_{SD} values. Therefore, the selection of appropriate QAA for different water types is necessary (Lee et al., 2015; Mao et al., 2018; Jiang et al., 2019; Uudeberg et al., 2019; Liu et al., 2019; Liu et al., 2020; Zeng et al., 2020; Bai et al., 2020).

However, only two water types (i.e., clear and turbid waters) were considered in the Jiang19 algorithm, and thus only two reference wavelengths (560 nm for clear waters and 754 nm for turbid water) were used to estimate $a(\lambda)$ and $b_{\mathrm{b}}(\lambda)$ at these two wavelengths. Although previous studies have confirmed the success of using 560 nm for clear waters (Lee et al., 2002) and 754 nm for highly turbid waters (Yang et al., 2013) as reference wavelengths in total absorption coefficient estimations, the assumption of a pure water absorption coefficient dominating the total absorption coefficient at these two wavelengths will probably be invalid in moderately turbid and extremely turbid waters. This is because that absorption coefficients of phytoplankton and non-algal particles at 560 and 754 nm will also be high in

moderately and extremely turbid waters (Jiang et al., 2021). Therefore, the applicability of the Jiang19 algorithm to different optical water types is still a challenge.

It is thus necessary to revisit the Jiang19 algorithm and improve its water type classification, reference wavelength selection, and water type corresponding QAA application to estimate more accurate Z_{SD} in waters with different optical properties.

1.4. Research Objectives of this Study

The following are the research objectives of this study that enabled addressing of the remaining problems identified in the Jiang19 algorithm to estimate more accurate Z_{SD} values.

(1) To develop a new algorithm for estimating more accurate Z_{SD} by selecting more appropriate QAA for different water types (Chapter II).

(2) To evaluate the performance of the developed Z_{SD} estimation algorithm using synthetic and *in situ* Z_{SD} values (Chapter III).

(3) To apply the developed Z_{SD} estimation algorithm to Medium Resolution Imaging Spectrometer (MERIS) data over Lake Kasumigaura between 2003 and 2012 to produce a long-term Z_{SD} product (Chapter IV).

Chapter II: Development of Z_{SD} estimation algorithm

2.1. Introduction

This chapter is about developing a Z_{SD} estimation algorithm. Since the applicability of Jiang19 algorithm to different optical water types is still a challenge as mentioned in chapter I, this chapter will detail the methodology and approach used to improve the Jiang19 algorithm hence achieving the objectives which are to: (1) develop a new algorithm for estimating Z_{SD} more accurately in waters with different optical properties by classifying more water types and by selecting a more appropriate reference wavelength and corresponding QAA for each water type; and (2) evaluate the performance of the developed Z_{SD} estimation algorithm by comparing it with the existing algorithm using synthetic data. Ultimately, the results achieved and discussions that are drawn from this study are given out.

2.2. Method

2.2.1. Synthetic data collection and generation

In this chapter, I collected one synthetic dataset from Jiang et al. (2021) which is called Synthetic Dataset I (N = 91,287) in Table 2.1. Synthetic Dataset I contains pairs of $R_{rs}(\lambda)$ spectra as well as total absorption coefficient ($a(\lambda)$) and total backscattering coefficient ($b_b(\lambda)$) values. The $R_{rs}(\lambda)$ spectra were generated from the simulated $a(\lambda)$ and $b_b(\lambda)$ values by using a bio-optical model proposed by Gordon et al. (1988) and Lee et al. (2002) having assumptions of a nadir viewing angle and optically deep waters. For Synthetic Dataset I, the chlorophyll-a concentration (C_{chl} (N = 63)), tripton concentration (C_{tr} (N = 63)), and CDOM

absorption coefficient at 440 nm (a_{CDOM} (440) (N = 23)) were varied with different intervals between 0.01–1000 mg/m^3, 0.01–1000 g/m^3, and 0.01–5 m^{-1}, respectively. More details on Synthetic Dataset I generation can be found in Jiang et al. (2021).

I then generated two other synthetic datasets to help with the development and evaluation of the algorithms. Synthetic Dataset II contains the diffuse attenuation coefficient ($K_{\mathrm{d}}(\lambda)$) values (N = 91287), which were generated from the simulated $a(\lambda)$ and $b_{\mathrm{b}}(\lambda)$ values in Synthetic Dataset I by using the semi-analytical equation (1.4) proposed by Lee et al. (2005, 2013).

Synthetic Dataset III consists of simulated Z_{SD} values (N = 91287). These Z_{SD} values were generated from the simulated $K_{\mathrm{d}}(\lambda)$ values (Synthetic Dataset II) and the corresponding $R_{\mathrm{rs}}(\lambda)$ values (Synthetic Dataset I) using the semi-analytical equation (1.5) developed by Jiang et al. (2019). A summary of all synthetic datasets is listed in Table 2.1.

Table 2.1 Summary of Synthetic Datasets I, II, and III

Synthetic Dataset	**I (Jiang et al., 2021)**	**II (This Study)**	**III (This Study)**	**Total Number of Data**	**Usage**
Parameter	$a(\lambda)$, $b_{\mathrm{b}}(\lambda)$, $R_{\mathrm{rs}}(\lambda)$	$K_{\mathrm{d}}(\lambda)$	Z_{SD}(0.01–44.68 m)	91287	Algorithm Validation

2.2.2. Development of a Z_{SD} estimation algorithm

Since previous studies have confirmed that semi-analytical Equations (1.4) and (1.5) are robust in waters with different optical properties if accurate $a(\lambda)$ and $b_{\mathrm{b}}(\lambda)$ are provided

(e.g., Lee et al., 2015; Jiang et al., 2019; Vundo et al., 2019; Lee et al., 2005; Yang et al., 2014), the efforts were focused on how to obtain more accurate $a(\lambda)$ and $b_{\mathrm{b}}(\lambda)$ values from $R_{\mathrm{rs}}(\lambda)$ spectra. A flowchart of the developed Z_{SD} estimation algorithm is shown in Figure 2.1. Comparing to the Jiang19 algorithm, four improvements were carried out in this study, which are detailed sequentially in the following paragraphs.

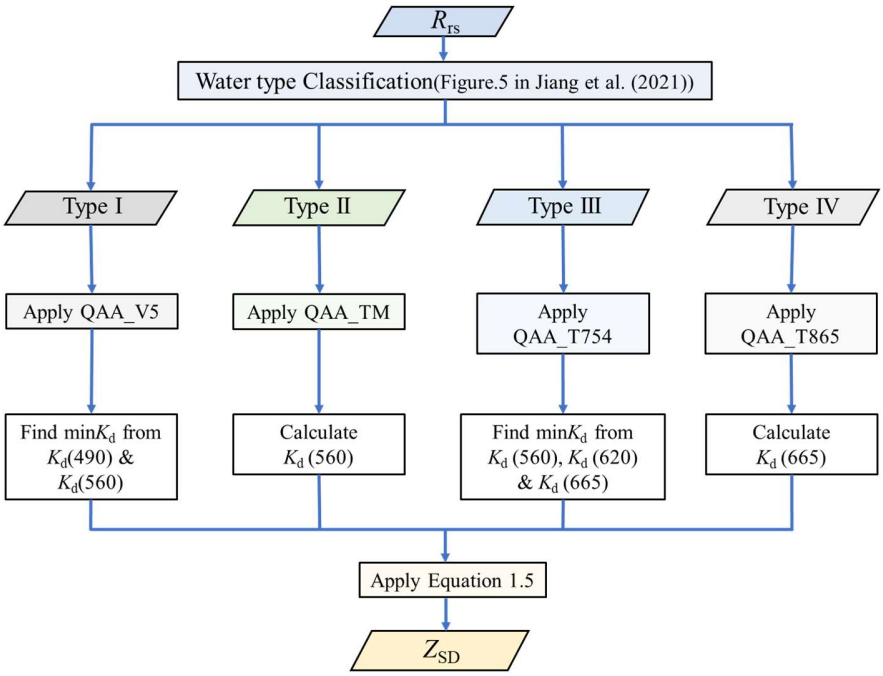


Figure 2.1. Flowchart of the Z_{SD} estimation algorithm proposed in this study

First, I adopted a recently proposed water type classification algorithm that can classify all waters into four types, clear (Type I), moderately turbid (Type II), highly turbid (Type III), and extremely turbid (Type IV), by comparing the values of R_{rs} at 490, 560, 620, and 754 nm (Jiang et al., 2021). This approach helped to more reasonably assume that the pure water absorption coefficient dominated the total absorption coefficients at the reference

wavelengths (i.e., 560 nm for Types I & II, 754 nm for Type III, and 865 nm for Type IV) (Lee et al., 2002; Jiang et al., 2021; Curtarelli et al., 2020). The water type classification algorithm is described as follows:

If $R_{rs}(490) > R_{rs}(560)$, then Type I waters,

Else, if $R_{rs}(490) > R_{rs}(620)$, then Type II waters,

Else, if $R_{rs}(754) > R_{rs}(490)$ and $R_{rs}(754) > 0.01$ sr^{-1}, then Type IV waters, (2.1)

Else, Type III waters.

Second, I selected a different reference wavelength and corresponding QAA for each water type to retrieve $a(\lambda_0)$ and $b_b(\lambda_0)$ values at the reference wavelength (λ_0). For Type I waters, I selected 560 nm as the reference wavelength and then used an empirical equation in QAA_V5 to estimate $a(560)$ due to its good performance in clear water (Lee et al., 2002; Joshi & D'Sa, 2018; Andrade et al., 2019; Deng et al., 2020). The equations are:

$$a(560) = a_w(560) + 10^{-1.146-1.366x-0.469x^2} \tag{2.2}$$

$$x = log_{10}\left(\frac{r_{rs}(443) + r_{rs}(490)}{r_{rs}(560) + 5\frac{r_{rs}(665)}{r_{rs}(490)}r_{rs}(665)}\right) \tag{2.3}$$

where r_{rs} is the remote sensing reflectance just below the water surface, $a_w(560)$ is the absorption coefficient of pure water at 560 nm, and $a(560)$ is the total absorption coefficient at 560 nm.

For Type II waters, I selected the same wavelength (i.e., 560 nm) as the reference wavelength but used a different empirical equation from QAA_TM (developed by Curtarelli et al. (2020)) to estimate $a(560)$. I adopted QAA_TM because of its good performance in moderately turbid waters compared to QAA_V6 developed by Lee et al. (2014) (Curtarelli et al., 2020). The equation is:

$$a(560) = a_{\mathrm{w}}(560) + 0.43\left(\frac{R_{\mathrm{rs}}(560)}{R_{\mathrm{rs}}(665) + R_{\mathrm{rs}}(709)}\right)^{-1.44} \tag{2.3}$$

where $R_{\mathrm{rs}}(560)$, $R_{\mathrm{rs}}(665)$, and $R_{\mathrm{rs}}(709)$ are the remote sensing reflectance just above the water surface at 560, 665, and 709 nm, respectively.

For Type III waters, 754 nm was selected as the reference wavelength, and the equations from QAA_T were adopted to estimate $a(754)$ (i.e., $a(754) \approx a_{\mathrm{w}}(754)$ (Yang et al., 2013)).

For Type IV waters, 865 nm was selected as the reference wavelength and the assumption of $a(865) \approx a_{\mathrm{w}}(865)$ was adopted, as suggested by Jiang et al. (2021). After estimation of $a(\lambda_0)$ at the reference wavelength (λ_0), the particulate backscattering coefficient ($b_{\mathrm{bp}}(\lambda_0)$) could be calculated by the following equation (Jiang et al., 2019; Lee et al., 2002):

$$b_{\mathrm{bp}}(\lambda_0) = \frac{\mu(\lambda_0)a(\lambda_0)}{1 - \mu(\lambda_0)} - b_{\mathrm{bw}}(\lambda_0) \tag{2.4}$$

Here, $b_{\mathrm{bw}}(\lambda_0)$ is the backscattering coefficient of pure water at the reference wavelength.

In the third effort, I employed two empirical models (for Types I and II, respectively) and one semi-analytical model (for both Types III and IV waters), which can represent reasonable shapes for particulate backscattering coefficients ($b_{\mathrm{bp}}(\lambda)$), thereby retrieving more accurate $a(\lambda)$ and $b_{\mathrm{b}}(\lambda)$ values at a given wavelength for each water type. For Type I waters, the spectral slope Y was calculated using an empirical equation in Lee et al. (2002):

$$Y = 2.0\left(1 - 1.2exp\left(-0.9\frac{r_{\mathrm{rs}}(443)}{r_{\mathrm{rs}}(560)}\right)\right) \tag{2.5}$$

For Type II waters, the spectral slope Y was calculated using a different empirical equation in Curtarelli et al. (2020):

$$Y = 0.5248exp\left(\frac{r_{\mathrm{rs}}(665)}{r_{\mathrm{rs}}(709)}\right) \tag{2.6}$$

For Types III and IV waters, the spectral slope Y was calculated using a semi-analytical equation in Yang et al. (2013):

$$Y = -372.99\left[log_{10}\left(\frac{\mu(754)}{\mu(779)}\right)\right]^2 + 37.286\,log_{10}\left(\frac{\mu(754)}{\mu(779)}\right) + 0.84 \tag{2.7}$$

Then $a(\lambda)$ and $b_{\mathrm{b}}(\lambda)$ at the given wavelength (λ) were calculated using the following equations (Lee et al., 2002):

$$b_{\mathrm{b}}(\lambda) = b_{\mathrm{bw}}(\lambda) + b_{\mathrm{bp}}(\lambda_0)\left(\frac{\lambda_0}{\lambda}\right)^Y \tag{2.8}$$

$$a(\lambda) = \frac{(1 - \mu(\lambda))b_{\mathrm{b}}(\lambda)}{\mu(\lambda)} \tag{2.9}$$

To summarize the second and third efforts, I selected QAA_V5 for Type I waters, QAA_TM for Type II waters, QAA_T with a reference wavelength of 754 nm (hereafter renamed QAA_T754) for Type III waters, and QAA_T with a reference wavelength of 865 nm (renamed QAA_T865) for Type IV waters.

In the fourth effort, I carried out a statistical analysis using Synthetic Dataset II to investigate the possibility of a minimum $K_{\mathrm{d}}(\lambda)$ for each water type. Figure 2.2 shows the number of synthetic water samples with the minimum $K_{\mathrm{d}}(\lambda)$ occurring at each visible wavelength for each water type. From the figure, the minimum $K_{\mathrm{d}}(\lambda)$ almost always occurred at 490 and 560 nm for Type I waters, 560 nm for Type II waters, 560, 620, and 665 nm for Type III waters, and at 665 nm for Type IV waters. Accordingly, I considered the relating wavelengths to find the minimum $K_{\mathrm{d}}(\lambda)$ for each water type. This effort avoided the selection of inappropriate wavelengths due to estimation errors in the previous steps (e.g., due to uncertainty in the estimations of *Y* values using equations (2.5)–(2.7).

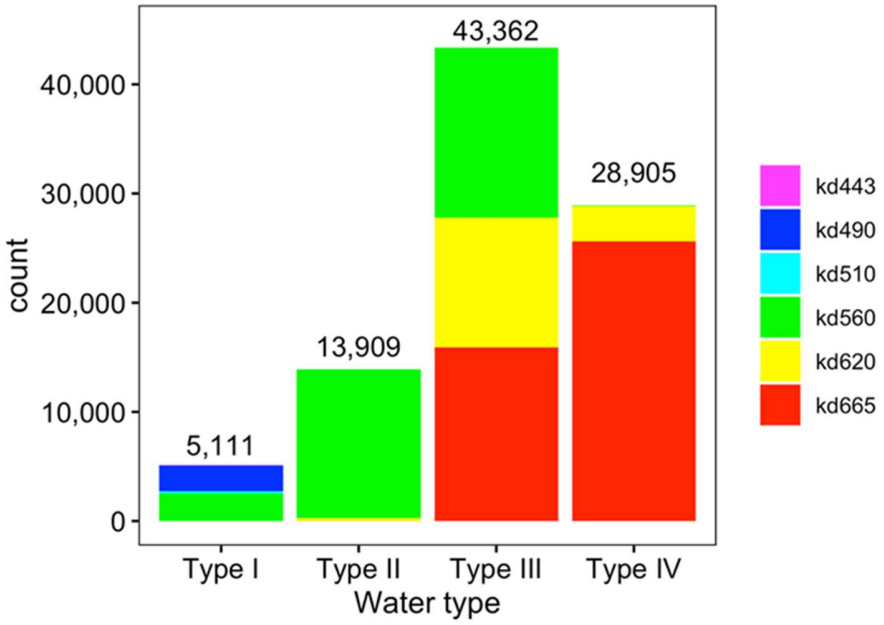


Figure 2.2. Statistical analysis using Synthetic Dataset II to identify the possibility of a minimum K_d (λ) for each water type

2.2.3. Accuracy assessment

I used the root mean square error (RMSE) in a log_{10} unit, the mean absolute percentage error (MAPE), bias, and the Nash-Sutcliffe efficiency (NSE) were used to evaluate the performance of the developed algorithm. The equations are as follows:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}\left[log_{10}(X_{\text{estimated},i}) - log_{10}(X_{\text{measured},i})\right]^2}{N}} \tag{2.10}$$

$$\text{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{X_{\text{estimated},i} - X_{\text{measured},i}}{X_{\text{measured},i}}\right| . 100\% \tag{2.11}$$

$$\text{Bias} = 10^{Y} - 1, \qquad Y = \frac{\sum_{i=1}^{N}\left[log_{10}\left(X_{\text{estimated},i}\right) - log_{10}\left(X_{\text{measured},i}\right)\right]}{N} \tag{2.12}$$

$$\text{NSE} = 1 - \frac{\sum_{i=1}^{N}\left(X_{\text{estimated},i} - X_{\text{measured},i}\right)^2}{\sum_{i=1}^{N}\left(X_{\text{measured},i} - \overline{X_{\text{measured}}}\right)^2} \tag{2.13}$$

where $X_{\text{estimated}}$ is the estimated Z_{SD} value, X_{measured} is the corresponding known Z_{SD} value, $\overline{X_{\text{measured}}}$ is the mean of known Z_{SD} values, and N is the number of data pairs. The regression results between the estimated and the known Z_{SD} values (i.e., R^2, slope, and intercept) were also used to help evaluate algorithm performance.

2.3. Results

2.3.1. Validation of Z_{SD}

Figure 2.3 compares known Z_{SD} values (i.e., Synthetic Dataset III) and estimated Z_{SD} values from Synthetic Dataset I using the Jiang19 algorithm and the algorithm developed in this study, respectively. From the figure, I observed many overestimations around the boundary between Type I and Type II waters as well as around the boundary between Type II and Type III waters when the Jiang19 algorithm was used (Figure 2.4(a)). These outliers resulted in very low values of R^2 (=0.008) and NSE (=−200.12), and high values of RMSE (0.30 in log_{10} Z_{SD} units), MAPE (116%), and bias (83%). In contrast, the outliers are well addressed in the new algorithm, with improved values of R^2 (0.97), NSE (0.90), RMSE (0.23

in log_{10} Z_{SD} units), and MAPE (65%) (Figure 2.3(b)). However, a systematic overestimation was found in the new algorithm (bias of 60%).

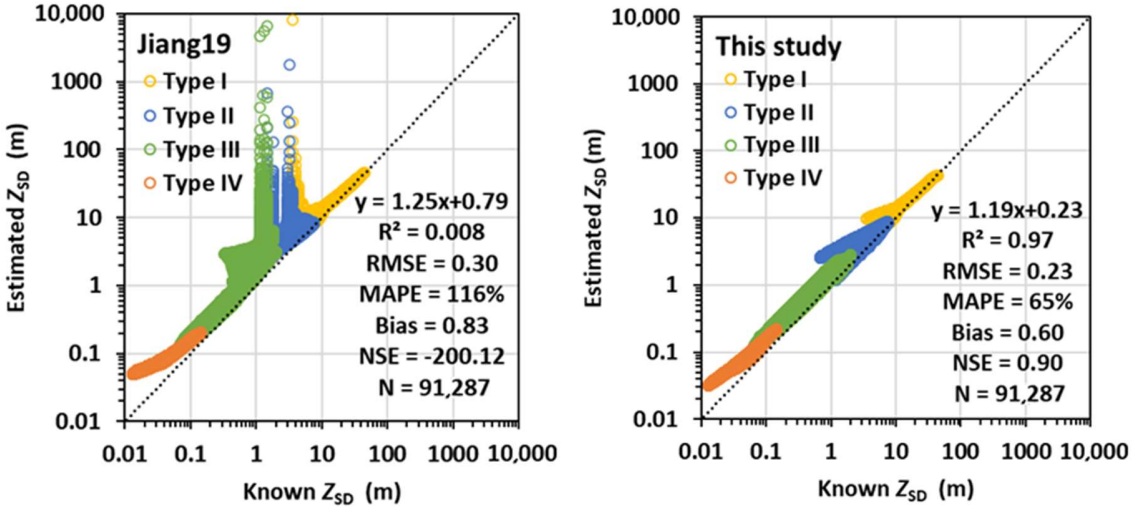
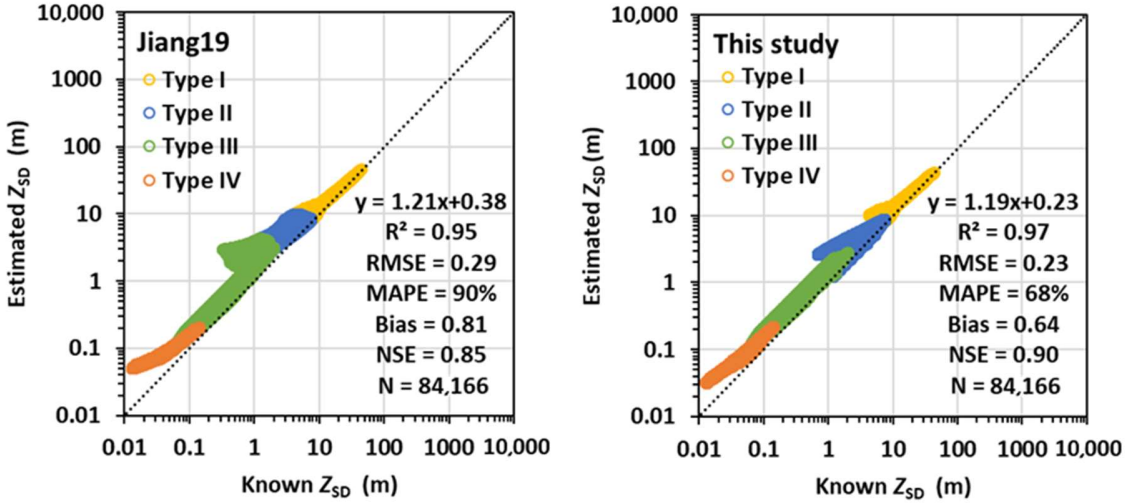


Figure 2.3. Comparison of known and estimated Z_{SD} values. The estimated Z_{SD} values were obtained from the simulated $R_{rs}(\lambda)$ in Synthetic Dataset I using the Jiang19 algorithm (a) and the new algorithm proposed in this study (b). The colors in the figures represent water types

2.3.2. Validation of Z_{SD} after removing outliers

Even when the outliers in Jiang et al.'s results were removed (N = 84,166), the proposed algorithm still performed better than the Jiang19 algorithm, with reduced values of RMSE (0.29 to 0.23 in log_{10} Z_{SD} units) and MAPE (90% to 68%) as well as increased values of R^2 (0.95 to 0.97) and NSE (0.85 to 0.90). Figure 2.4 compares known Z_{SD} values (i.e., Synthetic Dataset III) and estimated Z_{SD} values from Synthetic Dataset I (N = 84,166) using the Jiang19 algorithm and the algorithm developed in this study, respectively.



(**a**) (**b**)

Figure 2.4. Comparison of known and estimated Z_{SD} values. The estimated Z_{SD} values were obtained from the simulated $R_{rs}(\lambda)$ in Synthetic Dataset I (after removing outliers, N = 84,166) using the Jiang19 algorithm (a) and the new algorithm proposed in this study (b). The colors in the figures represent water types.

2.3.3. Z_{SD} Range for each water type

It was found that for water Type I, Z_{SD} values ranged from 5.54 m to 43.45 m. For water Type II, Z_{SD} values ranged from 1.19 m to 8.75 m. For water Type III, Z_{SD} values ranged from 0.12 m to 2.76 m and for water Type IV, Z_{SD} values ranged from 0.03 m to 0.22 m.

2.4. Discussion

2.4.1. Overestimation of Z_{SD} values

I developed a new algorithm to estimate Z_{SD} from Synthetic dataset I. Since this algorithm is based on several existing semi-analytical algorithms (Jiang et al., 2019; Lee et al., 2002; Yang et al., 2013; Jiang et al., 2021; Curtarelli et al., 2020), it can be applied to different optical water types, from clear to extremely turbid. The results from the synthetic dataset I (with a wide range of inherent optical properties, Z_{SD} values ranging from 0.01 to 44.68 m) provide strong evidence to confirm the applicability of the developed Z_{SD} estimation algorithm.

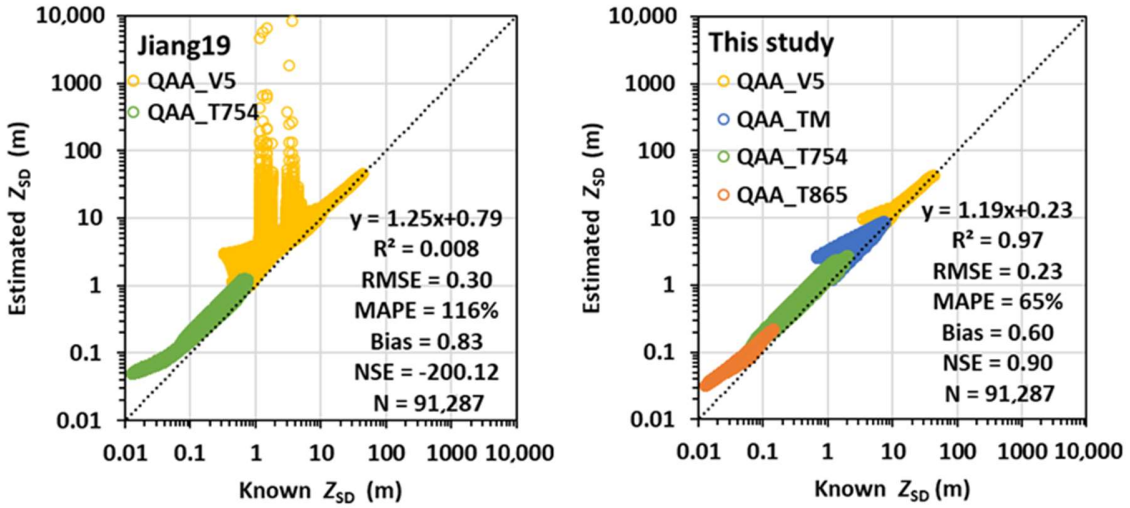
However, a systematic overestimation of Z_{SD} values was observed from the results when using synthetic dataset (Figure 2.3(b)). This is because the assumption for $a(\lambda)$ at the reference wavelengths resulted in an underestimation of $a(\lambda)$ values and thus an underestimation of $b_b(\lambda)$ values at the same wavelength. These underestimations propagated to the final estimated Z_{SD} values.

2.4.2. Influence of water type classification, QAA and reference wavelength selection on the accuracy of Z_{SD} estimation

Optical water type classification is very important because it helps to both clarify relationships between different properties inside a certain class and quantify the variation between classes (Uudeberg et al., 2019; Reinart et al., 2003). Recently, interest has grown in the application of the optical water type classification in the remote sensing of ocean color

(e.g., (Jiang et al., 2019; Jiang et al., 2021; Moore et al., 2014; Matsushita et al., 2015; Spyrakos et al., 2018; Balasubramanian et al., 2020)). In this study, I found that water type classification is a guide to select the most appropriate reference wavelength and corresponding QAA for the more accurate estimation of $a(\lambda)$ and $b_{\mathrm{b}}(\lambda)$ in waters with different optical properties. Therefore, I adopted the latest classification algorithm to classify all waters into four types. In contrast, the Jiang19 algorithm classified waters into only two types (clear and turbid) and used two corresponding QAAs (QAA_V5 and QAA_T754), thus generating large errors around boundaries of different water types (Figure 2.3(a)).

For example, in Figure 2.5(a), the Jiang19 algorithm classified some data samples as the clear water type and used QAA_V5 to estimate $a(\lambda)$ and $b_{\mathrm{b}}(\lambda)$; thus, large errors occurred in the Z_{SD} estimations (yellow points in Figure 2.5(a)). In contrast, I classified these data samples as Water Type II (moderately turbid waters), whereby QAA_TM was used for $a(\lambda)$ and $b_{\mathrm{b}}(\lambda)$ estimations and in turn improved the Z_{SD} estimations (blue points in Figure 2.5(b)).



(**a**) (**b**)

Figure 2.5. Comparison of known and estimated Z_{SD} values. The estimated Z_{SD} values were obtained from the simulated $R_{rs}(\lambda)$ in Synthetic Dataset I using the Jiang19 algorithm (a,) and the new algorithm proposed in this study (b). The colors in the figures represent the QAA used.

In addition, Jiang et al. (2021) suggested the use of QAA_V6 to estimate $a(\lambda)$ and $b_b(\lambda)$ at the reference wavelength of 665 nm for Type II waters. However, several previous studies have reported that QAA_V6 often failed in turbid waters (Rodrigues et al., 2017; Jiang et al., 2019; Yang et al., 2013, Curtarelli et al., 2020). Therefore, I used QAA_TM instead of QAA_V6 in this study. The results also show that QAA_TM performed better than QAA_V6 for Type II waters as shown in Figure 2.6.

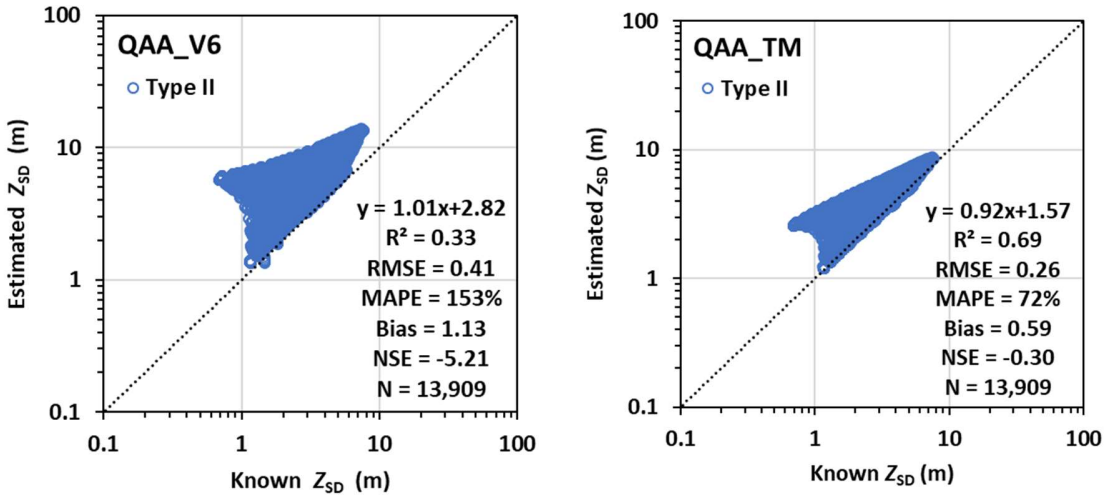


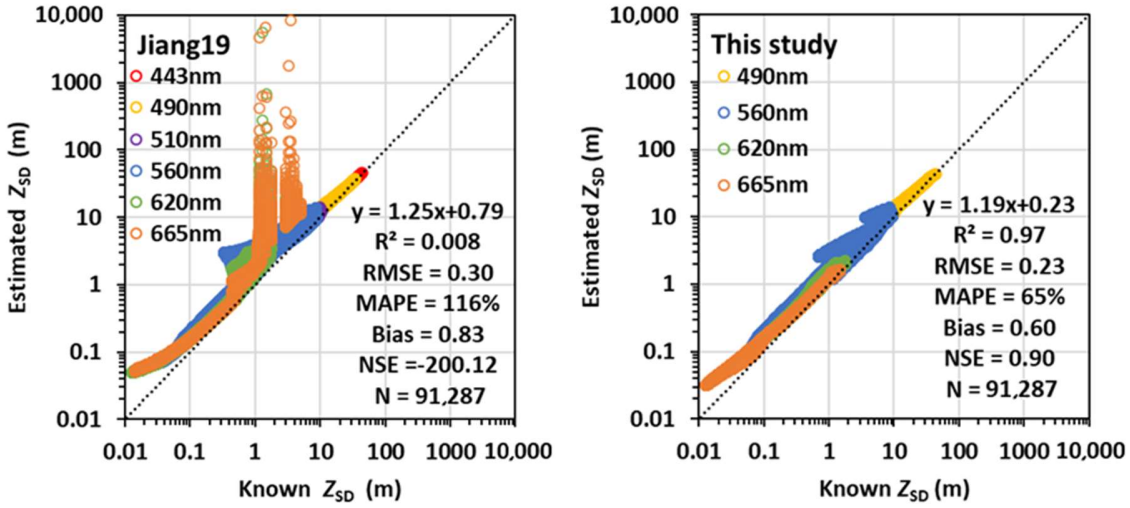
Figure 2.6. Comparison of known and estimated Z_{SD} values. The estimated Z_{SD} values were obtained from the simulated $R_{rs}(\lambda)$ in Synthetic Dataset I for water Type II only using the QAA_V6 algorithm (a) and QAA_TM (b,).

2.4.3. Influence of minimum K_d selection on the accuracy of Z_{SD} estimation (Water type I)

Overall, I improved the accuracy of $a(\lambda)$ and $b_b(\lambda)$ estimation by classifying waters into four types (effort 1), and then adopted different QAA to estimate $a(\lambda)$ and $b_b(\lambda)$ for each water type (efforts 2 and 3). These three efforts addressed most of the outliers caused by using the Jiang19 algorithm. However, some outliers remained after the above three efforts. This is because 620 or 665 nm was sometimes identified as the wavelength with the minimum K_d even for some Type I and II waters due to uncertainty in the Y estimation model

(Yang et al., 2013). The identification of 620 or 665 nm is obviously inappropriate as the wavelength with the minimum K_d in Type I and II waters should be at 490 or 560 nm based on the statistical results in Figure 2.2.

Therefore, in this study, I also limited the wavelengths used to estimate the minimum $K_d(\lambda)$ in each water type. This effort can make the wavelengths for estimating minimum $K_d(\lambda)$ close to the reference wavelength in each water type and thus can reduce errors due to uncertainty in the *Y* estimation model as seen in Figure 2.7(a,b).



(**a**) (**b**)

Figure 2.7. Comparison of known and estimated Z_{SD} values. The estimated Z_{SD} values were obtained from the simulated $R_{rs}(\lambda)$ in Synthetic Dataset I using the Jiang19 algorithm (a) and the new algorithm proposed in this study (b). The colors in the figures represent the wavelengths with the minimum $K_d(\lambda)$

2.5. **Conclusion**

In this chapter, a new semi analytical algorithm for Z_{SD} estimation was developed. The original Jiang19 algorithm showed systematic overestimations of Z_{SD} values but managed to be improved by 17% in this study. The improvement in Z_{SD} estimation was due to water type classification, and appropriate algorithm (QAA) selection according to water type classification. The developed Z_{SD} algorithm improved Jiang19 algorithm with an increased R^2 from 0.008 to 0.97, NSE from -200.12 to 0.90 and reduced RMSE from 0.30 to 0.23 (in $log_{10}\ Z_{SD}$ units), MAPE from 116% to 65%, and bias from 83% to 60% for synthetic dataset I. The developed Z_{SD} algorithm can be used to estimate Z_{SD} in different optical water types.

Chapter III: Validation of the developed Z_{SD} estimation algorithm

3.1. Introduction

After developing a new semi-analytical algorithm for Z_{SD} estimation in chapter II and check its performance by comparing it with the Jiang19 algorithm using synthetic dataset I, in this chapter the main objective was to validate the developed Z_{SD} estimation algorithm using an additional synthetic dataset (IV) as well as *in situ* dataset from 21 waters in Japan. This was done to further check the effectiveness and applicability of the developed Z_{SD} estimation algorithm in different optical water types.

3.2. Method

3.2.1. Synthetic data collection and generation

In this chapter, I collected another synthetic dataset from Jiang et al. (2021), which is called Synthetic Dataset IV (N = 1000) in Table 3.1. Synthetic Dataset IV contains pairs of $R_{rs}(\lambda)$ spectra as well as total absorption coefficient ($a(\lambda)$) and total backscattering coefficient ($b_b(\lambda)$) values. The procedures that this dataset was generated is the same as that of Synthetic Dataset I and as described in chapter II.

Unlike Synthetic Dataset I, for Synthetic Dataset IV, five ranges for chlorophyll-a concentration (C_{chl}), tripton concentration (C_{tr}), and CDOM absorption coefficient at 440 nm (a_{CDOM} (440)) values were randomly selected to generate 200 R_{rs} spectra. Therefore a total of 1000 R_{rs} spectra with C_{chl} values in the range of 0.01–1000 mg/ m^3, C_{tr} values in the

range of 0.01–1000 g/m^3, and a_{CDOM} (440) values in the range of 0.01–5 m^{-1} was obtained. More details on Synthetic Dataset IV generation can also be found in Jiang et al. (2021).

Thereafter I generated two other Synthetic Datasets namely Synthetic Dataset V and Synthetic Dataset VI. Synthetic Dataset V contains the diffuse attenuation coefficient ($K_{\mathrm{d}}(\lambda)$) values (N=1000) and Synthetic Dataset VI consists of simulated Z_{SD} values (1000). These $K_{\mathrm{d}}(\lambda)$ and Z_{SD} values were generated in the same manner as the Synthetic Dataset II and III in chapter II section 2.2.1. Synthetic Dataset IV, V and VI was used to validate the developed Z_{SD} estimation algorithm in chapter III and a summary of all Synthetic Datasets collected and generated in this chapter is listed in Table 3.1.

Table 3.1 Summary of Synthetic Datasets IV, V, and VI

Synthetic Dataset	**IV (Jiang et al., 2021)**	**V (This Study)**	**VI (This Study)**	**Total Number of Data**	**Usage**
Parameter	$a(\lambda)$, $b_{\mathrm{b}}(\lambda)$, $R_{\mathrm{rs}}(\lambda)$	$K_{\mathrm{d}}(\lambda)$	Z_{SD}(0.01–34.87 m)	1000	Algorithm Validation

3.2.2. *In situ* data collection

I compiled 305 *in situ* measured hyperspectral remote sensing reflectance ($R_{\mathrm{rs}}(\lambda)$) and Z_{SD} data pairs from 20 lakes and Tokyo Bay in Japan, which were collected between 2005 and 2020. Figure 3.1, hereafter shows the location of the sampling location (20 lake and Tokyo bay in Japan). These lakes include Lakes Kasumigaura, Shirakaba, Suwa, Hibara, Kugushi, Suigetsu, Megami, Suga, Mikata, Senbako, Tateshina Unagi, Saiko, Biwa, Inawashiro, Ikeda, Motosu, Yunoko, Akan and Syoji.

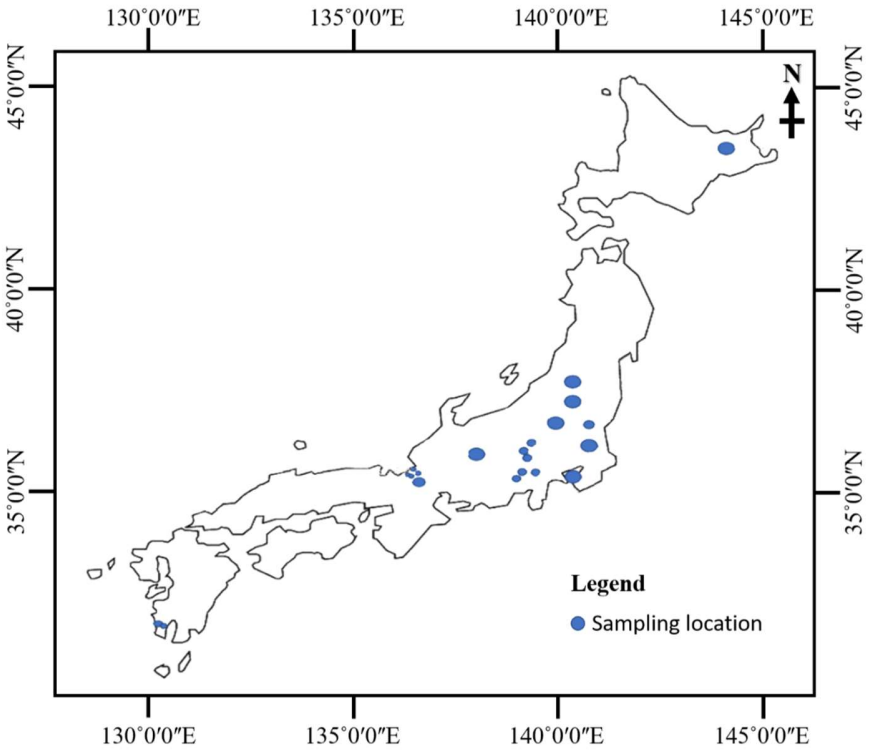


Figure 3.1. Study areas: Locations of 20 Japanese lakes and Tokyo Bay

The Z_{SD} values were measured with a 30 cm diameter white Secchi disk. The Z_{SD} data ranged from 0.3 to 16.4 m and covered clear to highly turbid waters. The *in situ* measured $R_{rs}(\lambda)$ values were obtained through the above-water approach. This approach measures the radiance of the skylight (L_s), the total upwelling radiance from the water (L_t), and the radiance from a standard gray board (L_g) using a FieldSpec® HandHeld spectroradiometer (ASD, Boulder, CO, USA) with a sensor zenith angle of 40° and an azimuth angle of 135° from the sun (Mobley, 1999). All measurements were carried out between 9:30 to 14:00 local time (three measurements were taken between 14:00 and 16:00). Then, the R_{rs} was calculated as follows:

$$R_{rs}=(L_t-\rho L_s)/\left(\frac{\pi}{R_g}L_g\right)-\Delta \quad (3.1)$$

where ρ is the water surface reflectance factor (0.028 when the wind speed is less than 5 m/s) (Mobley, 1999), R_g is the reflectance of the gray board, and Δ is the residual reflected skylight calculated using a method proposed by Jiang et al. (2020). All $R_{rs}(\lambda)$ spectra were then converted to MERIS bands based on the MERIS spectral response functions. This dataset was used to evaluate the performance of the developed Z_{SD} estimation algorithm.

3.2.3. Z_{SD} estimation

The developed Z_{SD} estimation algorithm developed in chapter II was applied to the Synthetic Dataset IV and *in situ* dataset to retrieve Z_{SD} values. However, in practical application some improvements were made to the Z_{SD} estimation algorithm developed in chapter II. These improvements were done for water Type II and Type III when using *in situ* dataset. The improvements are:

① Since equation (2.3) used a longer wavelength (709 nm) than those used in equation (2.2), and considering the effects of a low signal-to-noise ratio in a practical application, I followed Lee et al.'s (Lee et al., 2014) suggestion to use QAA_V5 for the waters with $R_{rs}(665) < 0.0015$ sr^{-1} (i.e., I still used equations (2.2) and (2.3)).

② Similar to the consideration for Type II waters, in Type III, I used a threshold of $R_{rs}(754) < 0.0015$ sr^{-1} and used QAA_TM for those waters with lower $R_{rs}(754)$. This approach also allowed avoiding the effects of a low signal-to-noise ratio in a practical application.

A flowchart of the developed Z_{SD} estimation algorithm improved in this chapter for practical application and being applied to the *in situ* dataset only is shown in Figure 3.2.

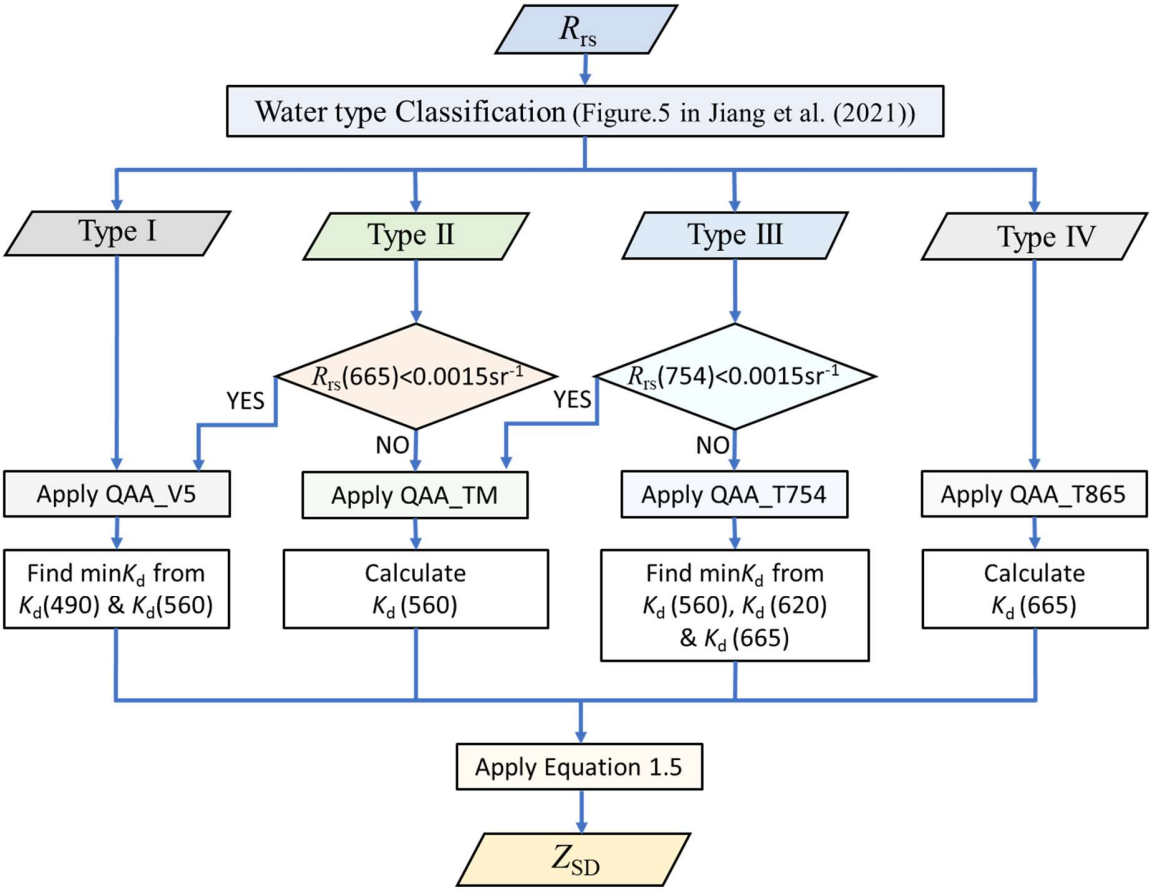


Figure 3.2. Flowchart of the Z_{SD} estimation algorithm applied to the *in situ* dataset.

3.2.4. Accuracy assessment

The root mean square error (RMSE) in a log_{10} unit, the mean absolute percentage error (MAPE), bias, and the Nash-Sutcliffe efficiency (NSE) were used to evaluate the performance of the developed algorithm using the following equations.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}\left[log_{10}(X_{\text{estimated},i}) - log_{10}(X_{\text{measured},i})\right]^2}{N}} \quad (3.2)$$

$$\text{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{X_{\text{estimated},i} - X_{\text{measured},i}}{X_{\text{measured},i}}\right| . 100\% \tag{3.3}$$

$$\text{Bias} = 10^{Y} - 1, \qquad Y = \frac{\sum_{i=1}^{N}\left[log_{10}\left(X_{\text{estimated},i}\right) - log_{10}\left(X_{\text{measured},i}\right)\right]}{N} \tag{3.4}$$

$$\text{NSE} = 1 - \frac{\sum_{i=1}^{N}\left(X_{\text{estimated},i} - X_{\text{measured},i}\right)^2}{\sum_{i=1}^{N}\left(X_{\text{measured},i} - \overline{X_{\text{measured}}}\right)^2} \tag{3.5}$$

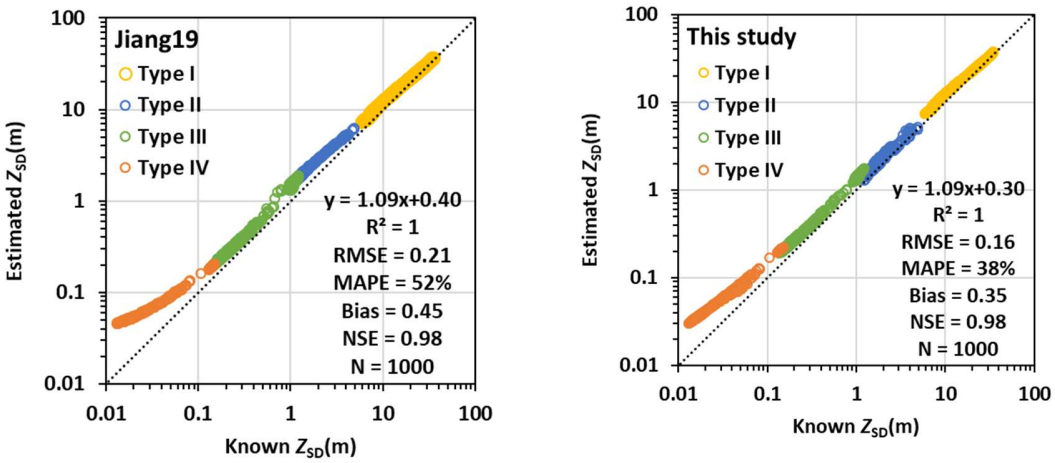
where $X_{\text{estimated}}$ is the estimated Z_{SD} value, X_{measured} is the corresponding *in situ* measured or known Z_{SD} value, $\overline{X_{\text{measured}}}$ is the mean of *in situ* measured or known Z_{SD} values, and N is the number of data pairs. The regression results between the estimated and the known or *in situ* measured Z_{SD} values (i.e., R^2, slope, and intercept) were also used to help evaluate algorithm performance.

3.3. Results

3.3.1. Validation of Z_{SD} using synthetic dataset IV

Figure 3.3 compares known Z_{SD} values (i.e., Synthetic Dataset VI) and estimated Z_{SD} values from Synthetic Dataset IV using the Jiang19 algorithm and the algorithm developed in this study, respectively. From the figure, overestimations for Type IV and around the boundary between Type II and Type III waters was still observed when the Jiang19 algorithm was used (Figure 3.3(a)). These systematic overestimations were improved in this study by

14% (Figure 3.3(b)). The results shows the algorithm developed in this study outperformed the Jiang19 algorithm with R^2 (=1), NSE (=0.98), RMSE (0.21 in log_{10} Z_{SD} units), MAPE (52%), and bias (45%) when using the Jiang19 algorithm versus R^2 (=1), NSE (=0.98), RMSE (0.16 in log_{10} Z_{SD} units), MAPE (38%), and bias (35%) when using the algorithm developed in this study.



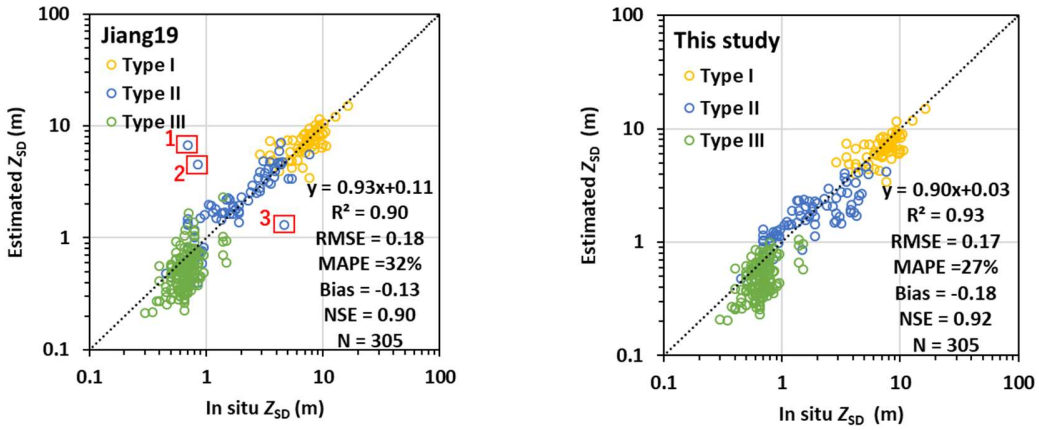
(a) **(b)**

Figure 3.3. Comparison of known and estimated Z_{SD} values. The estimated Z_{SD} values were obtained from the simulated $R_{rs}(\lambda)$ in Synthetic Dataset IV using the Jiang19 algorithm (a) and the new algorithm proposed in this study (b). The colors in the figures represent water types.

3.3.2. Validation of Z_{SD} using *in situ* dataset

Figure 3.4 compares the *in situ* measured and estimated Z_{SD} values. The estimated Z_{SD} values were obtained from *in situ* measured $R_{rs}(\lambda)$ spectra using the Jiang19 algorithm or the algorithm developed in this study. The results show that the new Z_{SD} estimation algorithm outperformed the Jiang19 algorithm with a reduced RMSE (from 0.18 to 0.17 in log_{10} unit) and MAPE (from 32% to 27%) as well as increased NSE (from 0.90 to 0.92) and R^2 values (from 0.90 to 0.93). In particular, I observed that three points with large errors due to the use of the Jiang19 algorithm (blue circles with red boxes in Figure 3.4(a)) were improved substantially by using the new algorithm (Figure 3.4(b)).

However, compared to the use of the Jiang19 algorithm, the bias due to the use of the new algorithm increased slightly, from a 13% underestimation to an 18% underestimation. In addition, unlike the results shown in Figures 2.3 and 3.3, systematic overestimation was not found in the application of *in situ* measured $R_{rs}(\lambda)$ spectra to estimate Z_{SD}, regardless of the use of the Jiang 19 algorithm or the new algorithm. Type IV waters were not found from *in situ* dataset.



(a) **(b)**

Figure 3.4. Comparison of *in situ* measured and estimated Z_{SD} values. The estimated Z_{SD} values were obtained from *in situ* measured $R_{rs}(\lambda)$ spectra using the Jiang19 algorithm (a) and the new algorithm proposed in this study (b). The colors in the figures represent water types.

3.4. Discussion

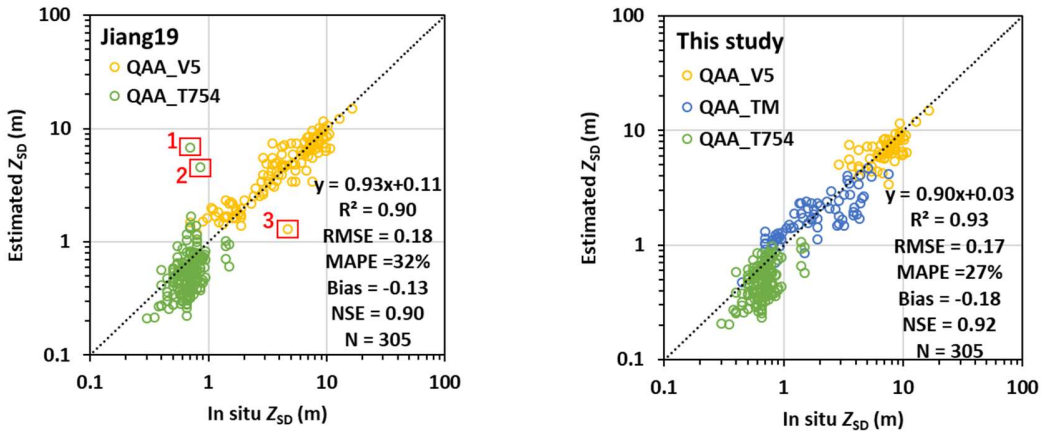
3.4.1. Overestimation of Z_{SD} values

Systematic overestimation was still observed when validating the Z_{SD} estimation algorithm developed in this study using Synthetic Dataset IV (Figure 3.3(b)) which were caused by the same factors described in chapter II. However, no overestimation of Z_{SD} values was observed when the developed algorithm was applied to the *in situ* data (Figure 3.4(b)). This is probably because the measurement errors contained in the *in situ* data hid (or offset) the systematic overestimation due to the mechanism of the developed algorithm. Therefore, further correction to mitigate the systematic overestimation was not carried out in this study.

3.4.2. The applicability of the developed Z_{SD} estimation

The developed Z_{SD} estimation algorithm in chapter II and its modification for practical application in this chapter has been applied to Synthetic Dataset IV and *in situ* dataset respectively and resulted reasonable Z_{SD} estimates for different optical water types (Figures 3.3(b) and 3.4(b)).

The four improvements (water type classification, QAA selection, reference wavelength selection, and the minimum K_d selection) made to the Jiang19 algorithm in chapter II has improved the accuracy of Z_{SD} values (Figure 3.3 and 3.4). For example, the Jiang19 algorithm classified the yellow point with red box 3 as clear water and two green points with red boxes 1 and 2 as turbid water, and thus used QAA_V5 and QAA_T754 to estimate $a(\lambda)$ and $b_b(\lambda)$, resulting finally in large errors in the Z_{SD} estimations (Figure 3.5(a)). These three Z_{SD} estimations were improved by classifying them as Type II waters and thus using QAA_TM to replace QAA_V5 and QAA_T754 (Figure 3.5(b)).



(a) **(b)**

Figure 3.5. Comparison of *in situ* measured and estimated Z_{SD} values. The estimated Z_{SD} values were obtained from *in situ* measured $R_{rs}(\lambda)$ spectra using the Jiang19 algorithm (a) and the new algorithm proposed in this study (b). The colors in the figures represent QAA used.

3.5. Conclusion

The developed Z_{SD} estimation algorithm has been applied to Synthetic Dataset IV and *in situ* dataset to estimate Z_{SD}. These datasets represent a wide range of water qualities (Z_{SD} ranges from 0.01 to 34.87 m for Synthetic Dataset IV and Z_{SD} ranges from 0.3 to 16.4 m for *in situ* dataset). The developed algorithm was able to improve the systematic overestimations of Z_{SD} values observed when using Synthetic Dataset IV by 10%. For Synthetic Dataset IV even though slightly overestimate was observed, yet the accuracy of all water type was improved with RMSE of 0.21 versus 0.16, MAPE of 52% versus 38%.

On the other hand, for *in situ* dataset, the accuracy was improved too, with RMSE of 0.18 versus 0.17, MAPE of 32% versus 27%. The results obtained from all datasets show that the four improvements described in Chapter II worked, and thus increased Z_{SD} estimation accuracy and thus the developed Z_{SD} estimation algorithm is expected to estimate more accurate Z_{SD} values in various types of waters.

Chapter IV: Application of the developed Z_{SD} estimation algorithm

4.1. Introduction

Information on water clarity obtained by field survey is still limited in terms of temporal frequency, spatial coverage, and representativeness even though Z_{SD} is one of the simplest measures of water properties. In contrast, earth observation data from satellite remote sensing can provide a feasible source to monitor long-term changes in Z_{SD} over large regions due to the revisit capacities and wide-swath observation of sensors onboard satellite platforms (Doron et al., 2011). To ensure management practices achieve sustainable development in the field of water environment, a routine monitoring by remote sensing technique is vital (Jiang et al., 2019; Vundo et al., 2019; Matsushita et al., 2012; Setiawan et al., 2019).

In chapter II the Z_{SD} estimation algorithm for estimating more accurate Z_{SD} values has been developed. Thereafter, the accuracy of the developed Z_{SD} estimation algorithm was validated using both Synthetic Dataset and *in situ* data to check its performance in chapter III.

Therefore, in this chapter the main objective is to apply the developed Z_{SD} estimation algorithm to MERIS satellite images of Lake Kasumigaura to obtain the time-series Z_{SD} product and analyze their changing trends.

4.2.Methods

4.2.1. Study area

The study area is Lake Kasumigaura (36.03° N, 140.40° E), which is the second biggest lake in Japan. It has an area of 167.6 km² (only west part) and an average water depth of 4 m, and is a typically shallow turbid lake. The Z_{SD} values of Lake Kasumigaura range from 0.05 m to 3.8 m (NIES, 2020). The chlorophyll-a concentrations range from 12.0 mg/m³ to 148.6 mg/m³ while the TSS range from 4.1 g/m³ to 45.6 g/m³ in this lake. In Lake Kasumigaura, there are 10 routine monitoring sites with a monitoring frequency of 1 month.

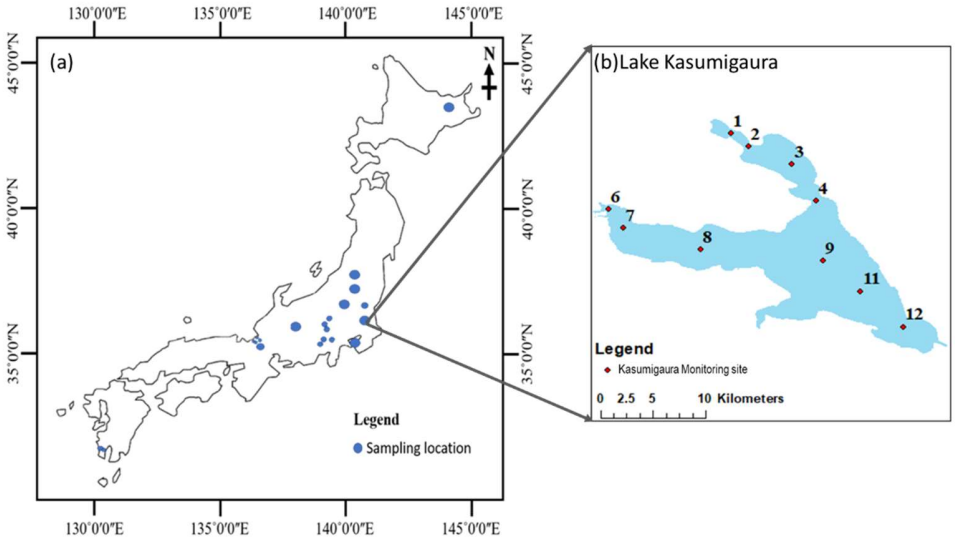


Figure 4.1. Study areas: (a) Location of Lake Kasumigaura, and (b) sampling sites of the Lake Kasumigaura Database.

4.2.2. *In situ* data collection

I acquired *in situ* measured Z_{SD} values at 10 stations in Lake Kasumigaura (Figure 4.1(b)) between January 2003 and March 2012 from the Lake Kasumigaura database, which were collected by the National Institute for Environmental Studies (NIES), Japan (NIES,2020). The Z_{SD} values ranged from 0.1 m to 1.8 m. This data was used for validation of the Z_{SD} values retrieved from satellite images by comparing the matchups (data acquired at the same day) and the long-term Z_{SD} time-series.

4.2.3. Satellite image processing

The MERIS Level-1B data between 2002 and 2012 for Lake Kasumigaura were collected from the European Space Agency (ESA, https://merisfrs-merci-ds.eo.esa.int, accessed on 04/02/2018). MERIS data was selected due to their good spatial (300 × 300 m) and spectral (15 bands in visible and near-infrared domains) characteristics. To remove the smile effect radiometric correction was done to the downloaded images after they have been clipped to the Lake Kasumigaura area.

Thereafter, atmospheric correction was performed using the Case-2 Regional Processor in BEAM Visualization and Analysis Tool version 5.0 (BEAM VISAT 5.0), and clouds and cloud shadows were identified using the IdePix module in Sentinel Application Platform (SNAP). I then masked out the pixels with clouds, cloud buffers, cloud shadows, and failed atmospheric correction. As a result, a total of 507 images were downloaded but only 200 images remained for the Z_{SD} estimation. The improved Z_{SD} estimation algorithm for

practical application in chapter III was applied to the remained MERIS images to retrieve Z_{SD} values for Lake Kasumigaura.

The correction of variations of visibility due to changes in the solar zenith angle was done to the MERIS derived Z_{SD} values using the Verschuur`s method (Verschuur, 1997). Then, I generated the monthly estimated Z_{SD} values by averaging all daily estimated Z_{SD} values in the same month. Due to the reason that, some sites in Lake Kasumigaura are very close to the shoreline hence being influenced by the land, a long-term Z_{SD} time-series at only seven sites (i.e. Sites 3, 4, 7, 8, 9, 11 and 12) was generated. The three remaining sites (i.e. Sites 1, 2 and 6) of Lake Kasumigaura were excluded from the comparison.

Finally, 19 matchups from Lake Kasumigaura were compiled by matching the acquisition times of *in situ* measured Z_{SD} values and satellite images (acquired on the same day). These matchups were used to further evaluate the performance of the developed Z_{SD} estimation algorithm. In addition, the averages of the Z_{SD} values derived from 3-by-3 pixels were used as satellite-estimated values.

4.2.4. Accuracy assessment

The root mean square error (RMSE) in a log_{10} unit, the mean absolute percentage error (MAPE), bias, and the Nash-Sutcliffe efficiency (NSE) were used to evaluate the performance of the developed algorithm. The equations are as follows:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}\left[log_{10}\left(X_{\mathrm{estimated},i}\right) - log_{10}\left(X_{\mathrm{measured},i}\right)\right]^2}{N}} \quad (4.1)$$

$$\text{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{X_{\text{estimated},i} - X_{\text{measured},i}}{X_{\text{measured},i}}\right| . 100\% \quad (4.2)$$

$$\text{Bias} = 10^{Y} - 1, \qquad Y = \frac{\sum_{i=1}^{N}\left[log_{10}\left(X_{\text{estimated},i}\right) - log_{10}\left(X_{\text{measured},i}\right)\right]}{N} \quad (4.3)$$

$$\text{NSE} = 1 - \frac{\sum_{i=1}^{N}\left(X_{\text{estimated},i} - X_{\text{measured},i}\right)^2}{\sum_{i=1}^{N}\left(X_{\text{measured},i} - \overline{X_{\text{measured}}}\right)^2} \quad (4.4)$$

where $X_{\text{estimated}}$ is the estimated Z_{SD} value from MERIS image, X_{measured} is the corresponding *in situ* measured Z_{SD} value, $\overline{X_{\text{measured}}}$ is the mean of *in situ* measured Z_{SD} values, and N is the number of data pairs. The regression results between the estimated and the *in situ* measured Z_{SD} values (i.e., R^2, slope, and intercept) were also calculated to evaluate algorithm performance.

4.3. Results

4.3.1. Validation using Z_{SD} matchups

Figure 4.2 compares the estimated Z_{SD} values from MERIS images using the newly developed algorithm to the *in situ* measured Z_{SD} values. The *in situ* measured Z_{SD} values were obtained from the Kasumigaura database. All matchups were classified to Type III waters based on their MERIS-derived $R_{\text{rs}}(\lambda)$ spectra. From the figure, the MERIS-derived Z_{SD} values were consistent with the *in situ* measured Z_{SD} values, with a MAPE value of 15%,

an RMSE value of 0.08 (in log_{10} unit), and a bias value of −9%. The R^2 value of 0.57 and the slope of 0.82 also indicate the good performance of the new Z_{SD} estimation algorithm. Since the Jiang19 algorithm also selected the same reference band (band 12 of MERIS) and the corresponding QAA (QAA_T754) for all matchups, identical results were obtained when we used the Jiang19 algorithm (data not shown). Systematic overestimation was not found in the application of MERIS data.

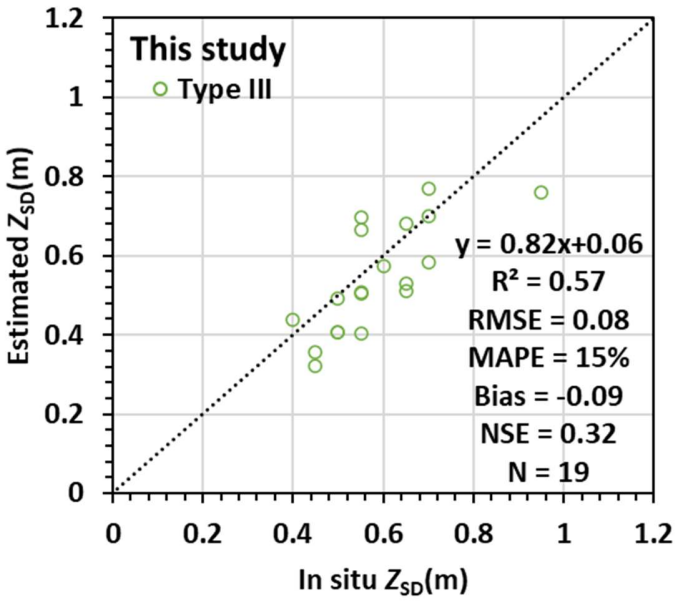
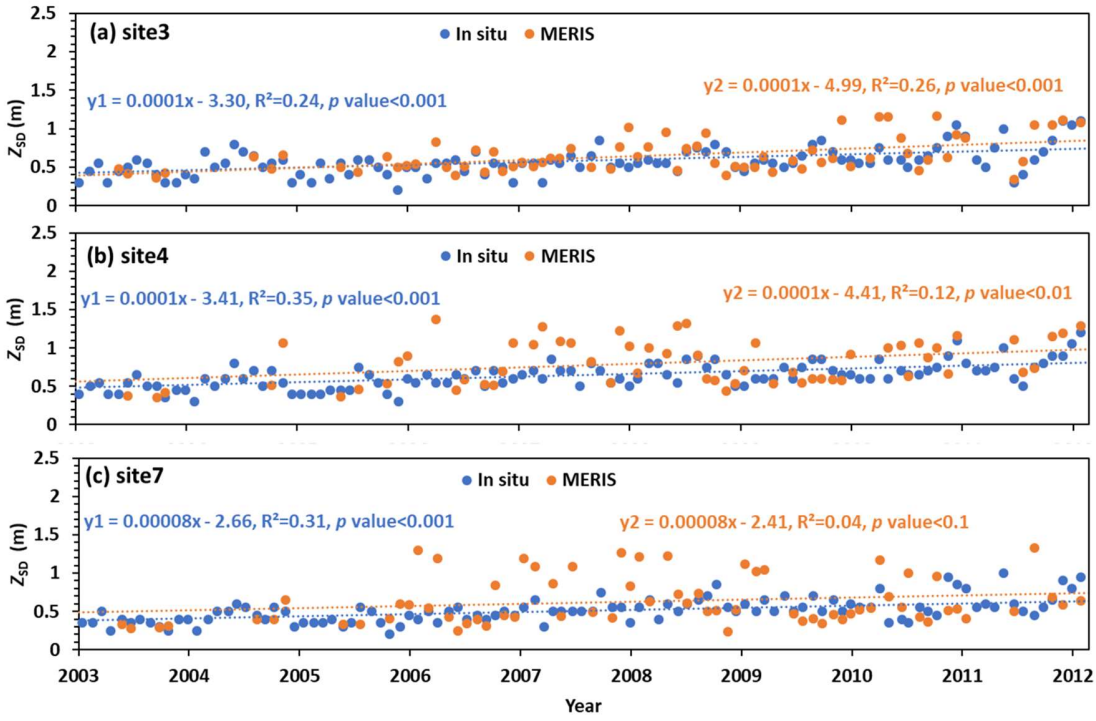


Figure 4.2. Comparisons of the *in situ* measured Z_{SD} values and the estimated Z_{SD} values from MERIS data using the new Z_{SD} estimation algorithm.

4.3.2. Long-term Z_{SD} product from MERIS Data

Figure 4.3 shows the comparison of the monthly *in situ* and estimated Z_{SD} values in Lake Kasumigaura for a long-term period of 2003-2012. The monthly *in situ* Z_{SD} values were obtained from NIES Kasumigara database while the monthly estimated Z_{SD} values were estimated using the developed Z_{SD} estimation algorithm from available MERIS data. From Figure 4.3, some estimated Z_{SD} values from satellite images (orange dots) matched well with the *in situ* measured Z_{SD} values (blue dots). In some months, the estimated Z_{SD} values were marked higher than their corresponding *in situ* measured Z_{SD} values.



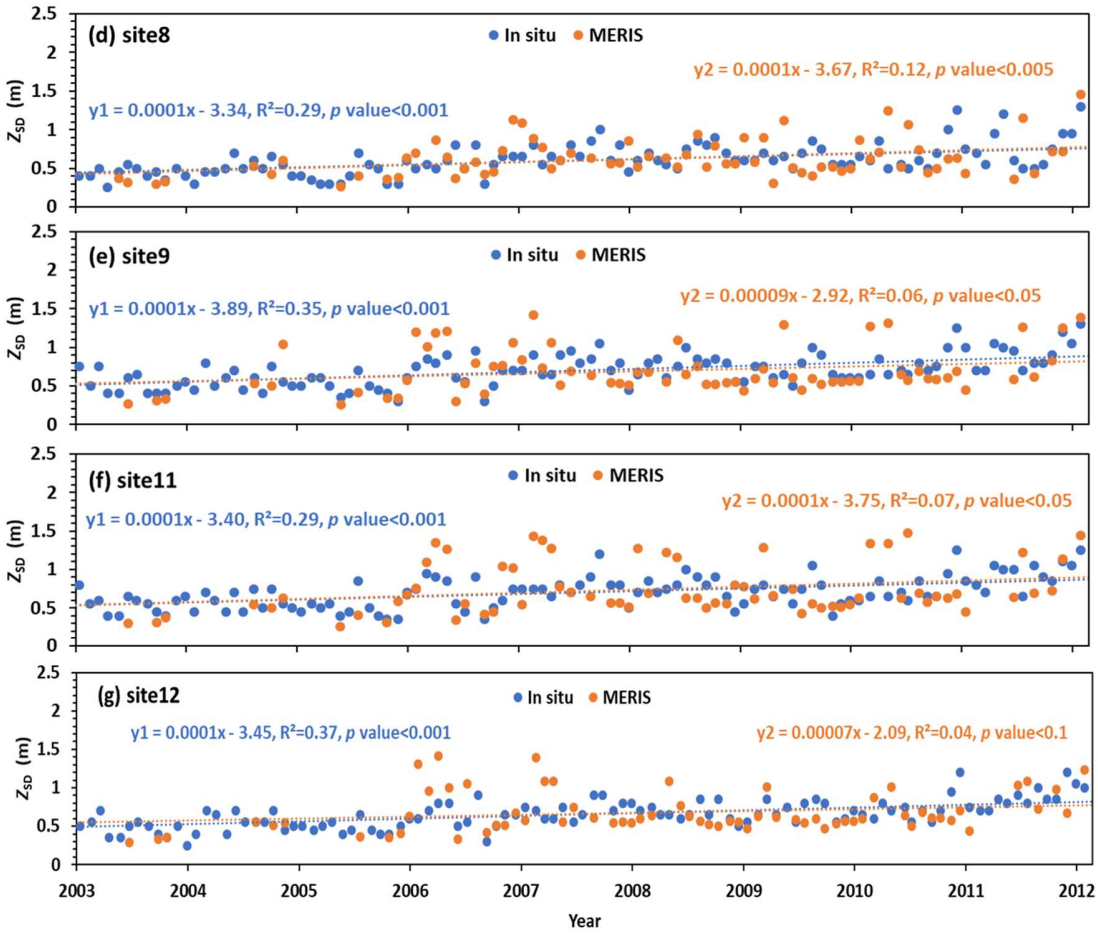


Figure 4.3. Long-term Z_{SD} time-series at seven sites in Lake Kasumigaura from 2003 to 2012. Blue dots represent *in situ* measured Z_{SD} values, Orange dots represent estimated Z_{SD} values from MERIS images using the developed Z_{SD} estimation algorithm.

Since Lake Kasumigaura belongs to water Type III according to the *in situ* R_{rs} spectra (Figure 4.4), therefore equations for estimating a and b_{bp} for water type III in the developed Z_{SD} estimation algorithm were used (i.e. QAA_T754 was used). This was done to correct the overestimation of Z_{SD} values observed in Figure 4.3.

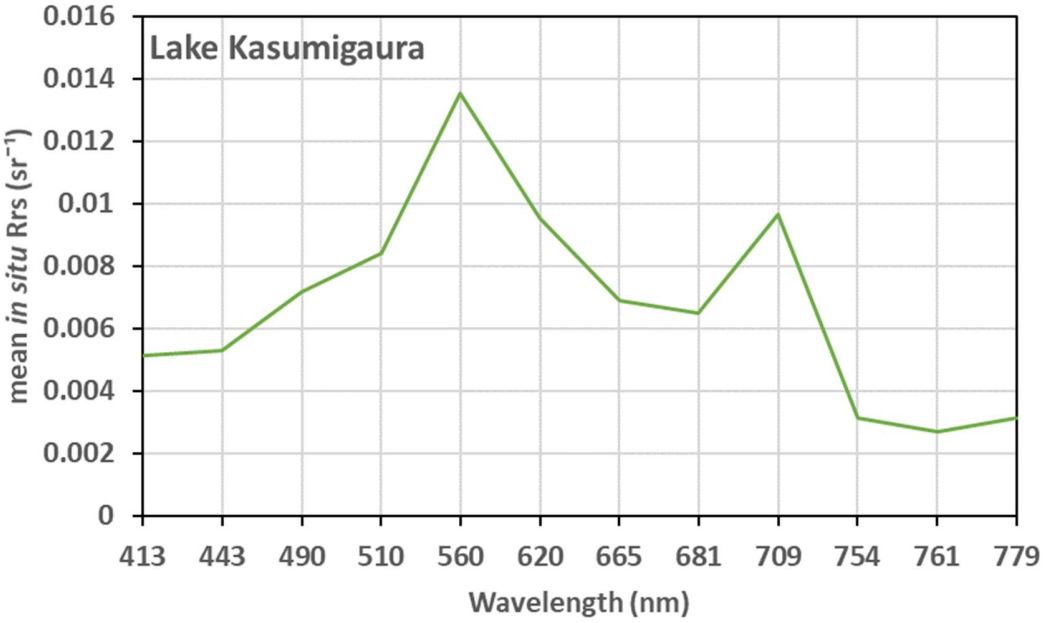


Figure 4.4. *In situ* R_{rs} spectra of Lake Kasumigaura.

Figure 4.5 shows the comparison of the monthly *in situ* and estimated Z_{SD} values in Lake Kasumigaura for a long-term period of 2003-2012. From Figure 4.4, estimated Z_{SD} values from satellite images (orange triangles) matched well with the *in situ* measured Z_{SD} values (blue dots). Both *in situ* and estimated Z_{SD} values in Lake Kasumigaura show comparable increasing trends from 2003 to 2012 with slope > 0 and p value <0.001.

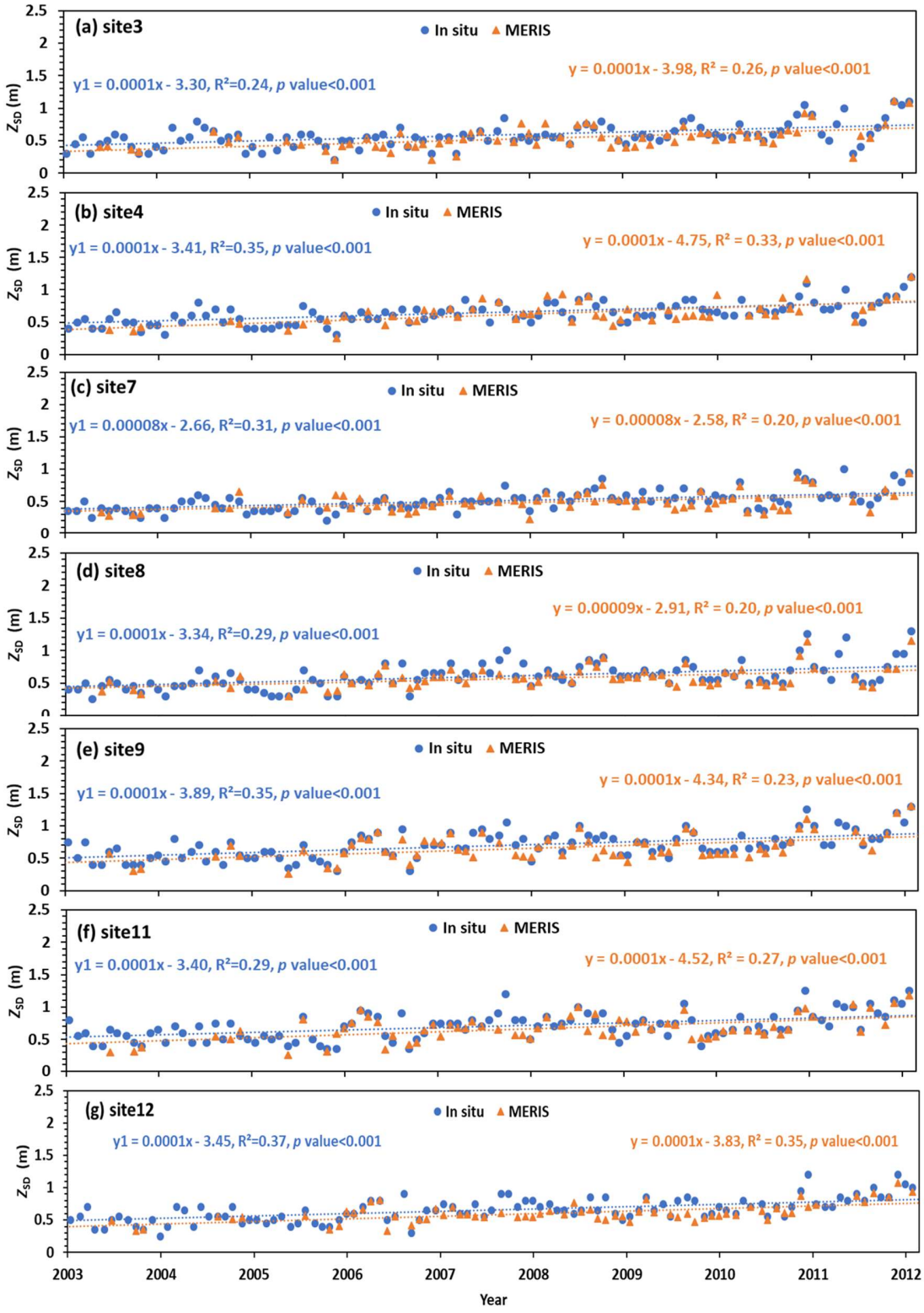

(a) site3
In situ
MERIS
y1 = 0.0001x - 3.30, R²=0.24, p value<0.001
y = 0.0001x - 3.98, R² = 0.26, p value<0.001
(b) site4
In situ
MERIS
y1 = 0.0001x - 3.41, R²=0.35, p value<0.001
y = 0.0001x - 4.75, R² = 0.33, p value<0.001
(c) site7
In situ
MERIS
y1 = 0.00008x - 2.66, R²=0.31, p value<0.001
y = 0.00008x - 2.58, R² = 0.20, p value<0.001
(d) site8
In situ
MERIS
y = 0.00009x - 2.91, R² = 0.20, p value<0.001
y1 = 0.0001x - 3.34, R²=0.29, p value<0.001
(e) site9
In situ
MERIS
y = 0.0001x - 4.34, R² = 0.23, p value<0.001
y1 = 0.0001x - 3.89, R²=0.35, p value<0.001
(f) site11
In situ
MERIS
y1 = 0.0001x - 3.40, R²=0.29, p value<0.001
y = 0.0001x - 4.52, R² = 0.27, p value<0.001
(g) site12
In situ
MERIS
y1 = 0.0001x - 3.45, R²=0.37, p value<0.001
y = 0.0001x - 3.83, R² = 0.35, p value<0.001
Z_SD (m)
0
0.5
1
1.5
2
2.5
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
Year

Figure 4.5. Long-term Z_{SD} time-series at seven sites in Lake Kasumigaura from 2003 to 2012. Blue dots represent *in situ* measured Z_{SD} values, Orange triangles represent estimated Z_{SD} values from MERIS images using the developed Z_{SD} estimation algorithm.

4.4. Discussion

4.4.1. The applicability of the developed Z_{SD} estimation on satellite data.

The accuracy of Z_{SD} values estimated from satellite images are affected by atmospheric correction (Mao et al., 2018; Jiang et al., 2019; Vundo et al., 2019; Setiawan et al., 2019). In this chapter atmospheric correction was performed using the Case-2 Regional Processor in BEAM VISAT 5.0. Case-2 Regional Processor was selected due to its better performance in Lake Kasumigaura (Jiang et al., 2019). The developed Z_{SD} estimation algorithm in Chapter II was applied to MERIS images from 2002 to 2012 of Lake Kasumigaura and reasonable results were obtained (Figure 4.2). Just like when using *in situ* data to estimate Z_{SD} values, no systematic overestimation of Z_{SD} values was observed when using satellite data (Z_{SD} matchups) as well. The reason is probably because the atmospheric correction and measurement errors contained in the satellite data offset the systematic overestimation due to the mechanism of the developed algorithm.

However, many points in Figure 4.3 were overestimated. This is because imperfect atmospheric correction affects water type classification hence wrong QAA is used to estimate a and b_b which in turn affects Z_{SD} estimation. For example some samples were classified as water Type I or Type II and used QAA_V5 or QAA_TM respectively to estimate a and b_b

in Figure 4.3 which resulted to overestimation. Despite the effect caused by atmospheric correction and measurement errors in Z_{SD} values estimated from satellite images, time variation and number of data between estimated Z_{SD} values and *in situ* measured Z_{SD} values within a month lead to the observed discrepancies (Mao et al., 2018; Wang et al., 2019). In addition, discrepancies between the measured and estimated Z_{SD} (i.e. overestimation, underestimation or scatter) in lake Kasumigaura are caused by surface scum and patches of cyanobacteria (Fukushima et al., 2016).

To address the overestimation observed in Figure 4.3, *in situ* R_{rs} spectra was used for water type classification (Figure 4.4). QAA_T754 was used to estimate a and b_b for all data samples hence estimate Z_{SD}values which in turn increased the R^2 and reduced the p value for all sites (Figure 4.5). Overall the estimated Z_{SD} matched well with the *in situ* measured Z_{SD}.

The results in this study are good evidence that by using the developed Z_{SD} estimation algorithm Z_{SD} values can be retrieved from satellite images. On the other hand further studies on atmospheric correction is vital to obtain more accurate Z_{SD} values from satellite images since according to Jiang et al. (2019) there is no any algorithm which can work effectively in various water types.

4.4.2. Usefulness of Z_{SD} estimation from satellite image

Recently remote sensing technique has simplified the process of large scale and long-term data acquisition in terms of time, cost and labor. By using satellite images, analyzing the changing trend of water quality parameters is one of the goal in the field of water quality

monitoring. Different studies have proved the applicability of satellite images in Z_{SD}values retrieval (Fukushima et al., 2016; Jiang et al., 2019; Vundo et al., 2019; Setiawan et al., 2019; Bai et al., 2020).

In this chapter, the long-term Z_{SD} time-series of Lake Kasumigaura was developed by applying the Z_{SD} estimation algorithm in chapter II on MERIS satellite images during 2002 to 2012. The goal was to analyze the changing trend of Z_{SD} values in Lake Kasumigaura. The obtained Z_{SD} changing trends agreed well with the trends of *in situ* data. The trend shows that, Z_{SD} is significant increasing from 2003 to 2012 (Figure 4.5).

By using the developed Z_{SD} estimation algorithm, long-term Z_{SD} time-series can be obtained from satellite images even from those waters which lack *in situ* data. This contribution is very important to water environment management sector where by Z_{SD} values can be directly retrieved from satellite images.

4.5. Conclusion

The developed Z_{SD} estimation algorithm was applied to MERIS satellite images to obtain the long-term Z_{SD} time-series for lake Kasumigaura. From the long-term Z_{SD} time-series, the estimated Z_{SD} value trends agreed well with *in situ* Z_{SD} trends. Likewise, the results from the 19 matchups showed that the estimated Z_{SD} in this study agreed well with the *in situ* Z_{SD} with RMSE of 0.08, MAPE of 15%, NSE of 0.32 and bias of -0.09. These results indicate that satellite data can be used for Z_{SD} monitoring.

Chapter V: General conclusions

A semi-analytical algorithm for estimating Z_{SD} values from remote sensing data has been developed in this study. The developed Z_{SD} estimation algorithm considers different optical properties in four water types, thus allowing the selection of the most optimal reference wavelength and QAA for more accurate estimation of total absorption and backscattering coefficients.

More accurate total absorption and backscattering coefficients lead to more accurate $K_d(\lambda)$ estimation, which in turn improved Z_{SD} estimations compared to the previous study. In addition, constraining the wavelength range of minimum $K_d(\lambda)$ in each water type contributed to the improvement of Z_{SD} estimation accuracy.

The developed Z_{SD} estimation algorithm was validated using two Synthetic Datasets and *in situ* dataset with a wide range of water qualities (Z_{SD} ranges from 0.01 to 44.68 m for synthetic data) in Chapter II, (Z_{SD} ranges from 0.01 to 34.87 m and 0.3 to 16.4 m for Synthetic and *in situ* data respectively) in Chapter III. The results showed that the developed algorithm is applicable to different optical water types with more accurate Z_{SD} estimations (MAPE improved from 116% to 65% when using Synthetic Dataset I, 52% to 38% when using Synthetic Dataset IV and from 32% to 27%. when using *in situ* dataset). In addition, the developed algorithm was able to correct some outliers that were obvious in Jiang19algorithm.

The developed Z_{SD} estimation algorithm was also applied to MERIS satellite images of Lake Kasumigaura to obtain the Z_{SD} values from 2003 to 2012. Validation results using matchups showed that the improved Z_{SD} algorithm can retrieve accurate Z_{SD} values. On the

other hand, changing trends of Z_{SD} obtained from satellite images in Lake Kasumigaura agreed well with the corresponding trends obtained from *in situ* Z_{SD} data during the study period. These results indicate that, the developed Z_{SD} estimation algorithm can be used to monitor Z_{SD} trends from satellite images.

References

Alikas, K., & Kratzer, S. (2017). Improved retrieval of Secchi depth for optically-complex waters using remote sensing data. *Ecological Indicators*, *77*, 218–227. https://doi.org/10.1016/j.ecolind.2017.02.007

Andrade, C., Alcântara, E., Bernardo, N., & Kampel, M. (2019). An assessment of semi-analytical models based on the absorption coefficient in retrieving the chlorophyll-a concentration from a reservoir. *Advances in Space Research*, *63*(7), 2175–2188. https://doi.org/10.1016/j.asr.2018.12.023

Bai, S., Gao, J., Sun, D., & Tian, M. (2020). Monitoring water transparency in shallow and eutrophic lake waters based on goci observations. *Remote Sensing*, *12*(1). https://doi.org/10.3390/RS12010163

Balasubramanian, S. V., Pahlevan, N., Smith, B., Binding, C., Schalles, J., Loisel, H., Gurlin, D., Greb, S., Alikas, K., Randla, M., Bunkei, M., Moses, W., Nguyễn, H., Lehmann, M. K., O'Donnell, D., Ondrusek, M., Han, T. H., Fichot, C. G., Moore, T., & Boss, E. (2020). Robust algorithm for estimating total suspended solids (TSS) in inland and nearshore coastal waters. *Remote Sensing of Environment*, *246*, 1–52. https://doi.org/10.1016/j.rse.2020.111768

Bowers, D. G., Roberts, E. M., Hoguane, A. M., Fall, K. A., Massey, G. M., & Friedrichs, C. T. (2020). Secchi Disk Measurements in Turbid Water. *Journal of Geophysical Research: Oceans*, *125*(5), 1–9. https://doi.org/10.1029/2020JC016172

Canfield, D., & Langeland, K. (1985). Relations between water transparency and maximum depth of macrophyte colonization in lakes. In *Journal of Aquatic Plant …* (Vol. 23, pp. 25–28).

Chang, N., Luo, L., Wang, X. C., Song, J., Han, J., & Ao, D. (2020). A novel index for assessing the water quality of urban landscape lakes based on water transparency. *Science of the Total Environment*, *735*(13), 139351. https://doi.org/10.1016/j.scitotenv.2020.139351

Curtarelli, V. P., Barbosa, C. C. F., Maciel, D. A., Flores Júnior, R., Carlos, F. M., Novo, E. M. L. de M., Curtarelli, M. P., & da Silva, E. F. F. (2020). Diffuse attenuation of clear water tropical reservoir: A remote sensing semi-analytical approach. *Remote Sensing*, *12*(17), 1–23. https://doi.org/10.3390/rs12172828

Deng, L., Zhou, W., Cao, W., Wang, G., Zheng, W., Xu, Z., Li, C., Yang, Y., Xu, W., Zeng, K., & Hu, S. (2020). Evaluating semi-analytical algorithms for estimating inherent optical properties in the South China Sea. *Optics Express*, *28*(9), 13155. https://doi.org/10.1364/oe.390859

Devlin, M. J., Barry, J., Mills, D. K., Gowen, R. J., Foden, J., Sivyer, D., & Tett, P. (2008). Relationships between suspended particulate material, light attenuation and Secchi

depth in UK marine waters. *Estuarine, Coastal and Shelf Science*, *79*(3), 429–439. https://doi.org/10.1016/j.ecss.2008.04.024

Doron, M., Babin, M., Hembise, O., Mangin, A., & Garnesson, P. (2011). Ocean transparency from space: Validation of algorithms estimating Secchi depth using MERIS, MODIS and SeaWiFS data. *Remote Sensing of Environment*, *115*(12), 2986–3001. https://doi.org/10.1016/j.rse.2011.05.019

Doron, M., Babin, M., Mangin, A., & Hembise, O. (2007). Estimation of light penetration, and horizontal and vertical visibility in oceanic and coastal waters from surface reflectance. *Journal of Geophysical Research: Oceans*, *112*(6), 1–15. https://doi.org/10.1029/2006JC004007

Duntley, S. Q. (1952). The visibility of submerged objects. *Visibility Lab., Mass. Inst. Tech. Scripps Institution of Oceanography, San Diego.*, 74.

Fleming-lehtinen, V., & Fleming-lehtinen, V. (2016). *Secchi depth in the Baltic Sea – an indicator of eutrophication* (Vol. 1, Issue November).

Fleming-Lehtinen, V., & Laamanen, M. (2012). Long-term changes in Secchi depth and the role of phytoplankton in explaining light attenuation in the Baltic Sea. *Estuarine, Coastal and Shelf Science*, *102–103*, 1–10. https://doi.org/10.1016/j.ecss.2012.02.015

Fukushima, T., Matsushita, B., Oyama, Y., Yoshimura, K., Yang, W., Terrel, M., Kawamura, S., & Takegahara, A. (2016). Semi-analytical prediction of Secchi depth using remote-sensing reflectance for lakes with a wide range of turbidity. *Hydrobiologia*, *780*(1), 5–20. https://doi.org/10.1007/s10750-015-2584-7

Gordon, R., Brown, O. B., Evans, H., Brown, W., Smith, C., Baker, K., & Clark, D. K. (1988). A Semianalytic Radiance Model of Ocean Color. *Journal of Geophysical Research*, *93*(8).

Gower, J., King, S., Borstad, G., & Brown, L. (2005). Detection of intense plankton blooms using the 709 nm band of the MERIS imaging spectrometer. *International Journal of Remote Sensing*, *26*(9), 2005–2012. https://doi.org/10.1080/01431160500075857

Holland, R. E. (1993). Changes in Planktonic Diatoms and Water Transparency in Hatchery Bay, Bass Island Area, Western Lake Erie Since the Establishment of the Zebra Mussel. *Journal of Great Lakes Research*, *19*(3), 617–624. https://doi.org/10.1016/S0380-1330(93)71245-9

Jiang, D., Matsushita, B., Lehmann, M. K., Moses, W. J., & Donnell, D. O. (2021). Remotely Estimating Total Suspended Solids Concentration in Various Waters Using a Novel Semi-Analytical Method. *Remote Sensing of Environment*, *258*(October 2020). https://doi.org/10.1016/j.rse.2021.112386

Jiang, D., Matsushita, B., Setiawan, F., & Vundo, A. (2019). An improved algorithm for estimating the Secchi disk depth from remote sensing data based on the new

underwater visibility theory. *ISPRS Journal of Photogrammetry and Remote Sensing*, *152*(April), 13–23. https://doi.org/10.1016/j.isprsjprs.2019.04.002

Jiang, D., Matsushita, B., & Yang, W. (2020). A simple and effective method for removing residual reflected skylight in above-water remote sensing reflectance measurements. *ISPRS Journal of Photogrammetry and Remote Sensing*, *165*(July 2019), 16–27. https://doi.org/10.1016/j.isprsjprs.2020.05.003

Joshi, I. D., & D'Sa, E. J. (2018). An estuarine-tuned quasi-analytical algorithm (QAA-V): Assessment and application to satellite estimates of SPM in Galveston Bay following Hurricane Harvey. *Biogeosciences*, *15*(13), 4065–4086. https://doi.org/10.5194/bg-15-4065-2018

Lee, G. F., Jones-Lee, A., & Rast, W. (1995). Secchi depth as a water quality parameter. *Publication Pending*.

Lee, Z., Carder, K. L., & Arnone, R. A. (2002). Deriving inherent optical properties from water color: a multiband quasi-analytical algorithm for optically deep waters. *Applied Optics*, *41*(27), 5755. https://doi.org/10.1364/ao.41.005755

Lee, Z., 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. https://doi.org/10.1364/ao.37.006329

Lee, Z., Carder, K. L., Mobley, C. D., Steward, R. G., & Patch, J. S. (1999). Hyperspectral remote sensing for shallow waters: 2 Deriving bottom depths and water properties by optimization. *Applied Optics*, *38*(18), 3831. https://doi.org/10.1364/ao.38.003831

Lee, Z., Hu, C., Shang, S., Du, K., Lewis, M., Arnone, R., & Brewin, R. (2013). Penetration of UV-visible solar radiation in the global oceans: Insights from ocean color remote sensing. *Journal of Geophysical Research: Oceans*, *118*(9), 4241–4255. https://doi.org/10.1002/jgrc.20308

Lee, Z. P., Du, K. P., & Arnone, R. (2005). A model for the diffuse attenuation coefficient of downwelling irradiance. *Journal of Geophysical Research C: Oceans*, *110*(2), 1–10. https://doi.org/10.1029/2004JC002275

Lee, Z. P., Shang, S., Hu, C., Du, K., Weidemann, A., Hou, W., Lin, J., & Lin, G. (2015). Secchi disk depth: A new theory and mechanistic model for underwater visibility. *Remote Sensing of Environment*, *169*, 139–149. https://doi.org/10.1016/j.rse.2015.08.002

Lee, Z., Shang, S., Du, K., & Wei, J. (2018). Resolving the long-standing puzzles about the observed Secchi depth relationships. *Limnology and Oceanography*, *63*(6), 2321–2336. https://doi.org/10.1002/lno.10940

Liu, D., Duan, H., Loiselle, S., Hu, C., Zhang, G., Li, J., Yang, H., Thompson, J. R., Cao, Z., Shen, M., Ma, R., Zhang, M., & Han, W. (2020). Observations of water

transparency in China's lakes from space. *International Journal of Applied Earth Observation and Geoinformation*, *92*(April), 102187. https://doi.org/10.1016/j.jag.2020.102187

Liu, X., Lee, Z., Zhang, Y., Lin, J., Shi, K., Zhou, Y., Qin, B., & Sun, Z. (2019). Remote sensing of secchi depth in highly turbid lake waters and its application with MERIS data. *Remote Sensing*, *11*(19). https://doi.org/10.3390/rs11192226

Liu, Y., Xiao, C., Li, J., Zhang, F., & Wang, S. (2020). Secchi disk depth estimation from China's new generation of GF-5 hyperspectral observations using a semi-analytical scheme. *Remote Sensing*, *12*(11). https://doi.org/10.3390/rs12111849

Mao, Y., Wang, S., Qiu, Z., Sun, D., & Bilal, M. (2018). Variations of transparency derived from GOCI in the Bohai Sea and the Yellow Sea. *Optics Express*, *26*(9), 12191. https://doi.org/10.1364/oe.26.012191

Matsushita, B., Yang, W., Chang, P., Yang, F., & Fukushima, T. (2012). A simple method for distinguishing global Case-1 and Case-2 waters using SeaWiFS measurements. *ISPRS Journal of Photogrammetry and Remote Sensing*, *69*, 74–87. https://doi.org/10.1016/j.isprsjprs.2012.02.008

Matsushita, B., Yang, W., Yu, G., Oyama, Y., Yoshimura, K., & Fukushima, T. (2015). A hybrid algorithm for estimating the chlorophyll-a concentration across different trophic states in Asian inland waters. *ISPRS Journal of Photogrammetry and Remote Sensing*, *102*, 28–37. https://doi.org/10.1016/j.isprsjprs.2014.12.022

McCullough, I. M., Loftin, C. S., & Sader, S. A. (2012). High-frequency remote monitoring of large lakes with MODIS 500m imagery. *Remote Sensing of Environment*, *124*(November 2011), 234–241. https://doi.org/10.1016/j.rse.2012.05.018

Mobley, C. D. (1999). Estimation of the remote-sensing reflectance from above-surface measurements. *Applied Optics*, *38*(36), 7442. https://doi.org/10.1364/ao.38.007442

Moore, T. S., Dowell, M. D., Bradt, S., & Ruiz Verdu, A. (2014). An optical water type framework for selecting and blending retrievals from bio-optical algorithms in lakes and coastal waters. In *Remote Sensing of Environment* (Vol. 143, pp. 97–111). https://doi.org/10.1016/j.rse.2013.11.021

Olmanson, L. G., Bauer, M. E., & Brezonik, P. L. (2008). A 20-year Landsat water clarity census of Minnesota's 10,000 lakes. *Remote Sensing of Environment*, *112*(11), 4086–4097. https://doi.org/10.1016/j.rse.2007.12.013

Reinart, A., Herlevi, A., Arst, H., & Sipelgas, L. (2003). Preliminary optical classification of lakes and coastal waters in Estonia and south Finland. *Journal of Sea Research*, *49*(4), 357–366. https://doi.org/10.1016/S1385-1101(03)00019-4

Rodrigues, T., Alcântara, E., Watanabe, F., & Imai, N. (2017). Retrieval of Secchi disk depth from a reservoir using a semi-analytical scheme. In *Remote Sensing of*

Environment (Vol. 198, pp. 213–228). https://doi.org/10.1016/j.rse.2017.06.018

Sandén, P., & Håkansson, B. (1996). Long-term trends in Secchi depth in the Baltic Sea. *Limnology and Oceanography*, *41*(2), 346–351. https://doi.org/10.4319/lo.1996.41.2.0346

Setiawan, F., Matsushita, B., Hamzah, R., Jiang, D., & Fukushima, T. (2019). Long-term change of the secchi disk depth in Lake Maninjau, Indonesia shown by landsat TM and ETM+ data. *Remote Sensing*, *11*(23), 1–20. https://doi.org/10.3390/rs11232875

Sòria-Perpinyà, X., Urrego, E. P., Pereira-Sandoval, M., Ruiz-Verdú, A., Soria, J. M., Delegido, J., Vicente, E., & Moreno, J. (2020). Monitoring water transparency of a hypertrophic lake (The albufera of València) using multitemporal sentinel-2 satellite images. *Limnetica*, *39*(1), 373–386. https://doi.org/10.23818/limn.39.24

Spyrakos, E., O'Donnell, R., Hunter, P. D., Miller, C., Scott, M., Simis, S. G. H., Neil, C., Barbosa, C. C. F., Binding, C. E., Bradt, S., Bresciani, M., Dall'Olmo, G., Giardino, C., Gitelson, A. A., Kutser, T., Li, L., Matsushita, B., Martinez-Vicente, V., Matthews, M. W., … Tyler, A. N. (2018). Optical types of inland and coastal waters. *Limnology and Oceanography*, *63*(2), 846–870. https://doi.org/10.1002/lno.10674

Swift, T. J., Perez-Losada, J., Schladow, S. G., Reuter, J. E., Jassby, A. D., & Goldman, C. R. (2006). Water clarity modeling in Lake Tahoe: Linking suspended matter characteristics to Secchi depth. *Aquatic Sciences*, *68*(1), 1–15. https://doi.org/10.1007/s00027-005-0798-x

Tyler, J. E. (1968). THE SECCHI DEPTH. *Limnology and Oceanography*, *XIII*(January).

Uudeberg, K., Ansko, I., Põru, G., Ansper, A., & Reinart, A. (2019). Using opticalwater types to monitor changes in optically complex inland and coastalwaters. *Remote Sensing*, *11*(19), 1–20. https://doi.org/10.3390/rs11192297

Verschuur, G. L. (1997). Transparency measurements in Garner Lake, Tennessee: The relationship between secchi depth and solar altitude and a suggestion for normalization of secchi depth data. *Lake and Reservoir Management*, *13*(2), 142–153. https://doi.org/10.1080/07438149709354305

Vundo, A., Matsushita, B., Jiang, D., Gondwe, M., Hamzah, R., Setiawan, F., & Fukushima, T. (2019). An overall evaluation of water transparency in Lake Malawi from MERIS data. *Remote Sensing*, *11*(3). https://doi.org/10.3390/rs11030279

Wang, F., Umehara, A., Nakai, S., & Nishijima, W. (2019). Distribution of region-specific background Secchi depth in Tokyo Bay and Ise Bay, Japan. *Ecological Indicators*, *98*(July 2018), 397–408. https://doi.org/10.1016/j.ecolind.2018.11.015

Wang, S., Lee, Z., Shang, S., Li, J., Zhang, B., & Lin, G. (2019). Deriving inherent optical properties from classical water color measurements: Forel-Ule index and Secchi disk depth. *Optics Express*, *27*(5), 7642. https://doi.org/10.1364/oe.27.007642

Wernand, M. R. (2010). On the history of the Secchi disc. *Journal of the European Optical Society*, *5*. https://doi.org/10.2971/jeos.2010.10013s

Wondie, A., Mengistu, S., Vijverberg, J., & Dejen, E. (2007). Seasonal variation in primary production of a large high altitude tropical lake (Lake Tana, Ethiopia): Effects of nutrient availability and water transparency. *Aquatic Ecology*, *41*(2), 195–207. https://doi.org/10.1007/s10452-007-9080-8

Yang, W., Matsushita, B., Chen, J., Yoshimura, K., & Fukushima, T. (2013). Retrieval of inherent optical properties for turbid inland waters from remote-sensing reflectance. *IEEE Transactions on Geoscience and Remote Sensing*, *51*(6), 3761–3773. https://doi.org/10.1109/TGRS.2012.2220147

Yang, W., Matsushita, B., Chen, J., Yoshimura, K., & Fukushima, T. (2014). Application of a semianalytical algorithm to remotely estimate diffuse attenuation coefficient in turbid inland waters. *IEEE Geoscience and Remote Sensing Letters*, *11*(6), 1046–1050. https://doi.org/10.1109/LGRS.2013.2284343

Zeng, S., Lei, S., Li, Y., Lyu, H., Xu, J., Dong, X., Wang, R., Yang, Z., & Li, J. (2020). Retrieval of secchi disk depth in Turbid Lakes from GOCI based on a new semi-analytical algorithm. *Remote Sensing*, *12*(9). https://doi.org/10.3390/RS12091516

Zhang, X., & Hu, L. (2009). Scattering by pure seawater at high salinity. *Optics Express*, *17*(15), 12685. https://doi.org/10.1364/oe.17.012685