Spatial Distribution of Diffuse Attenuation of Photosynthetic Active Radiation and Its Main Regulating Factors in Inland Waters of Northeast China

Jianhang Ma 1,2, Kaishan Song 1,*, Zhidan Wen 1, Ying Zhao 1,2, Yingxin Shang 1,2, Chong Fang 1,2 and Jia Du 1

1 Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences (CAS), Changchun 130102, China; mmjjhh105@sina.com (J.M.); wenzhidan@iga.ac.cn (Z.W.); zhaoying477@163.com (Y.Z.); goodlucksyx27@163.com (Y.S.); fangchong1991@gmail.com (C.F.); jiaqidu@iga.ac.cn (J.D.)
2 University of Chinese Academy of Sciences, Beijing 100049, China
* Correspondence: songks@iga.ac.cn; Tel.: +86-431-8554-2364

Academic Editors: Linhai Li, Claudia Giardino, Yunlin Zhang, Deepak R. Mishra, Xiaofeng Li and Prasad S. Thenkabail
Received: 15 August 2016; Accepted: 16 November 2016; Published: 21 November 2016

Abstract: Light availability in lakes or reservoirs is affected by optically active components (OACs) in the water. Light plays a key role in the distribution of phytoplankton and hydrophytes, thus, is a good indicator of the trophic state of an aquatic system. Diffuse attenuation of photosynthetic active radiation (PAR) (Kd(PAR)) is commonly used to quantitatively assess the light availability. The PAR and the concentration of OACs were measured at 206 sites, which covered 26 lakes and reservoirs in Northeast China. The spatial distribution of Kd(PAR) was depicted and its association with the OACs was assessed by grey incidences (GIs) and linear regression analysis. Kd(PAR) varied from 0.45 to 15.04 m$^{-1}$. This investigation revealed that reservoirs in the east part of Northeast China were clear with small Kd(PAR) values, while lakes located in plain areas, where the source of total suspended matter (TSM) varied, displayed high Kd(PAR) values. The GIs and linear regression analysis indicated that the TSM was the dominant factor in determining Kd(PAR) values and best correlated with Kd(PAR) ($R^2 = 0.906$, RMSE = 0.709). Most importantly, we have demonstrated that the TSM concentration is a reliable measurement for the estimation of the Kd(PAR) as 74% of the data produced a relative error (RE) of less than 0.4 in a leave-one-out cross validation (LOO-CV) analysis. Spatial transferability assessment of the model also revealed that TSM performed well as a determining factor of the Kd(PAR) for the majority of the lakes. However, a few exceptions were identified where the optically regulating dominant factors were chlorophyll-a (Chl-a) and/or the chromophoric dissolved organic matter (CDOM). These extreme cases represent lakes with exceptionally clear waters.

Keywords: light attenuation coefficient; optically active constituents; Northeast China; total suspended matter; water transparency

1. Introduction

Diffuse attenuation of photosynthetic active radiation (PAR), expressed as Kd(PAR), indicates the ability of solar radiation to penetrate a water column. The distribution of algae and hydrophytes, which contribute greatly to the lake’s primary production, is mainly influenced by the availability of light as well as other factors, for example, temperature and nutrition [1,2]. Euphotic zone depth ($z_{eu}$), an important input parameter for ecological models that estimate primary production of inland
waters [3,4], defined as the depth where 1% of the PAR just beneath the water surface remains [5], can be calculated using the equation \( z_{cd} = 4.6/K_d(\text{PAR}) \) [5] provided that a water column is homogeneous. The \( K_d(\text{PAR}) \) is also an important variable to research the heat transfer of lakes [6]. Global climate change and anthropogenic impacts have strong effects on a lakes' ecosystem and this may be depicted by changes in \( K_d(\text{PAR}) \). Therefore, research of \( K_d(\text{PAR}) \) and further understanding of its effects can significantly contribute to the development of new approaches for the management and protection of lake environments [7]. Although water transparency, measured with Secchi disk depth (SDD) by human eyes, can also represent the light properties of lakes [8], \( K_d(\text{PAR}) \) provides a more objective depiction as it is measured with advanced electro-optical instruments [9].

\( K_d(\text{PAR}) \) can be obtained by fitting the profile of PAR values measured at different depths of the water versus the corresponding depths, according to Lambert–Beer’s law [10–12]. Although portable instruments, such as the Li-cor 191, are capable for accurate measurements to estimate \( K_d(\text{PAR}) \), there are limitations in their application on large scale regions [12] as frequent in situ sampling is costly, labor-intensive and time-consuming. Optical remote sensing imagery is a cost-efficient method to obtain \( K_d(\text{PAR}) \) values at large regional scale due to the correlation between \( K_d(\text{PAR}) \) and water leaving radiance, and also due to its spatial and temporal resolution. Numerous remote sensing data such as Landsat/TM/OLI [13], Sea-viewing Wide Field-of-view Sensor (SeaWiFS) [14], the Moderate Resolution Imaging Spectroradiometer (MODIS) [15], the Medium Resolution Imaging Spectrometer (MERIS) [12,16,17] and the Geostationary Ocean Color Imager (GOCI) [18] have been applied to retrieve \( K_d(\text{PAR}) \) or \( K_d(490) \) (which is often used as an agent of \( K_d(\text{PAR}) \)).

Semi-analytical models [15,19–22] for \( K_d(\text{PAR}) \) or \( K_d(490) \) inversion emphasized the importance of inherent optical properties (IOPs) and improved the accuracy of the \( K_d(\text{PAR}) \) estimation in both open ocean waters (case-I) and coastal or inland waters (case-II). However, large uncertainties still existed in this type of algorithms for turbid and optically complex inland waters [15,19,21]. Though the applications may be confined to specified regions or seasons, empirical models have been widely used to derive \( K_d(\text{PAR}) \) from remote sensing data for both case-I and case-II waters [12,23–25]. The models were built by calibrating in situ \( K_d(\text{PAR}) \) with remote sensing reflectance at blue-green [14,24], red [12] or Near-infrared (NIR) [26] bands so the \( K_d(\text{PAR}) \) can be directly mapped from remote sensing images. \( K_d(\text{PAR}) \) is governed by the properties of natural water that include both dissolved and particulate organic as well as the inorganic material [2,5,27]. Therefore, the \( K_d(\text{PAR}) \) can be expressed as a function of the dominant one or some optically active components (OACs) whose concentrations can be estimated from remote sensing data empirically or semi-analytically. Thus, \( K_d(\text{PAR}) \) can be mapped indirectly from remote sensing images. Chl-a plays a significant role in optical property of case-I water so it is rational to estimate \( K_d(\text{PAR}) \) by the concentration of Chl-a that derived from satellite images [9,28]. Dominant factor of case-II water’s optical property varied dramatically from Chl-a [27] to TSM [10,29] or CDOM [1,30]. Sometimes, there were multiple dominant factors [31] and they changed by seasons [32]. Thereby, comprehensive analysis of the relationships between \( K_d(\text{PAR}) \) and OACs with in situ data was necessary before indirectly deriving \( K_d(\text{PAR}) \) from remote sensing data for inland waters. Models to estimate \( K_d(\text{PAR}) \) from OACs can be built with in situ data [33–35]. Factors that play a significant role in the variance of \( K_d(\text{PAR}) \) can be identified through the analysis [2] and further used as a guidance for the policy making in the protection of limnology environments.

Northeast China lake zone, with 882 lakes whose area was greater than 1 km² in 2010 [36], is one of the five lake zones of China [37]. Some of the lakes and reservoirs in Northeast China are featured with high TSM and CDOM due to the strong wind and shallow water depth, combined with rich soil organic matter that supply much terrigenous dissolved organic matter (DOM) into waters [38–41]. Thus, the optical property and their influence on \( K_d(\text{PAR}) \) may be different from other lake zones of China. Remote sensing may be the most suitable method to monitor environmental parameters for the widely distributed lakes. Investigations on \( K_d(\text{PAR}) \) in lakes of East China zone have been carried out [2,7,12,26,31]. Comparatively, much less research on \( K_d(\text{PAR}) \) in Northeast China lakes have been conducted. Thus, the analysis of dominate OACs of \( K_d(\text{PAR}) \) could provide a guidance for future
investigating of \( K_d(PAR) \) by remote sensing. The objectives of this paper are: (1) to depict the spatial distribution pattern of \( K_d(PAR) \) of lakes and reservoirs, which represents the trophic state of waters in Northeast China well; (2) to determine the dominant OACs of \( K_d(PAR) \) in waters of Northeast China using data collected from 26 lakes and reservoirs with three types of grey incidences (GIs); and (3) to build a relationship between \( K_d(PAR) \) and OACs that is meaningful for indirectly mapping \( K_d(PAR) \) from remotely sensed imagery.

2. Materials and Methodologies

2.1. Study Area

Northeast China is an important base of agriculture, forestry, energy and heavy industry. The region which extends from 38° N to 54.0° N and 116° E to 136° E, includes all of the Heilongjiang, Jilin and Liaoning Provinces, and parts of Inner Mongolia Province. Its topography is characterized by mountains to the east, north and west that surround the Sanjiang, Songnen, and Liaohe Plains (Figure 1a). The region has a temperate continental monsoon climate which is controlled by the East Asian monsoons. This climate is characterized for its cool and short summers and for its cold and long winters with the lakes being frozen in winters. The annual average temperature ranges between −4 and 12 °C with a gradual increase from north to south. Precipitation is generally higher in the summer and autumn seasons and decreases from 1100 mm in the southeast to 250 mm in the west [42].

Forests are predominantly distributed in mountainous areas where the soil erosion is weak. Grasslands are mainly distributed in the Inner Mongolia plateau where the precipitation is generally less than 400 mm (Figure 1b). A large amount of lakes and reservoirs with various features are distributed in Northeast China. Reservoirs are mainly located in mountainous areas and are very deep with long-narrow shapes. Unlike reservoirs, most of the lakes are generally shallow and are located in plain areas, particularly in Western Songnen Plain. Due to the character of environmental factors like unevenly distributed precipitation, high evaporation and geomorphology of terminal-flow areas, many fresh and saline water bodies are distributed in the Songnen Plain [39].

![Figure 1. Study area map: (a) Digital elevation model (DEM) and sampling lakes with corresponding lake ID; and (b) type of vegetation and spatial distribution of the \( K_d(PAR) \) derived from in situ measurements, combined with isotherm and isohyets in Northeast China.](image)

2.2. In Situ Data Collection and Analysis

In total, 206 stations with measurement of PAR collected in six field experiments were used. The stations covered 26 lakes and reservoirs in Northeast China. Details about the distribution of the
sample points are shown in Figure S1 (Supplementary Materials). There are only two lakes with areas less than 10 km$^2$. The water approximately 0.5 m below the surface was collected in the acid-washed HDPE bottles for laboratory analysis. The location of each station was recorded with a UniStrong G3 GPS uint. Water transparency (Secchi disk depth, SDD) was measured with a 30 cm diameter black and white quadrant (Secchi) disk. PAR was measured by the Li-Cor 193SA underwater spherical quantum sensor on the sunny side of the boat to avoid any shadow effects. After posing the sensor at one depth in the water, PAR value was continuously recorded for 15 s and output an averaged value by the data logger. This value was regarded as the PAR value at this depth. The PAR measurements were taken at no less than five point’s depth for each station. $K_d$(PAR) was determined by applying the exponential regression model which utilizes Equation (1) [33], provided that the water column was homogeneous. The results were accepted only if the coefficient of determination ($R^2$) was no less than 0.95 [7] and the number of depth points was no less than 4. In Equation (1), the PAR(Z) represents the PAR value at depth Z and PAR(0$^-$) represents the PAR value just beneath the surface of the water.

$$\text{PAR(Z)} = \text{PAR(0$^-$)} \times \exp\left[-K_d(\text{PAR}) \times Z\right] \quad (1)$$

2.3. Water Quality Parameters

To determine the water quality, we calculated the concentrations of TSM and Chl-a, and the absorption of CDOM, phytoplankton and non-algal particles (NAP), as follows.

TSM concentration: For all samples, the concentration of TSM was determined gravimetrically. Whatman GF/F glass fiber filters (47 mm in diameter, 0.7 µm in average pore size) were initially combusted at 400 °C for 4 h to remove any organic matters on the filters. After cooling, they were weighed before proceeding to filtration. The volumes of water samples to be filtered were determined by their turbidity. The used filters were stored at 4 °C and re-weighed after drying for 4 h at 105 °C. The concentration of the TSM was calculated by dividing the difference of weight by the volume of the corresponding water sample.

Chl-a concentration: Chl-a was determined spectrophotometrically [43,44]. A certain volume (V) of water sample was filtered through GF/F cellulose acetate membrane filters with 47 mm in diameter and 0.47 µm in pore size. The filters were frozen at −20 °C and stored under dark conditions until further analysis. Pigments were extracted by soaking the mashed filters in 10 mL of 90% acetone solution for 24 h under dark conditions. The supernatant was collected by centrifugation (5000 r/min, 20 min) and its absorbance at 630, 647, 664 and 750 nm was measured by the Shimadzu UV-2600 PC spectrophotometer. Concentration of Chl-a was calculated by Equation (2) [45], where OD(630), OD(647), OD(664) and OD(750) represented the absorbance at 630, 647, 664 and 750 nm, respectively. The number 10 is the volume of the acetone solution. V is the volume of water sample in liter. L is the cuvette path length in cm. The cuvette with path length of 1 cm was used in this study.

$$\text{Chl-a} = \{11.85 \times [\text{OD(664)} - \text{OD(750)}] - 1.54 \times [\text{OD(647)} - \text{OD(750)}] - 0.08 \times [\text{OD(630)} - \text{OD(750)}]\} \times 10 / (V \times L) \quad (2)$$

CDOM absorption: The generally high concentration of particles in inland waters results in the difficulties in collecting enough filtrate for measurement by solely filtering water samples through 0.22 µm filters as the particles block the pore easily. Thus, water samples were initially filtered through 0.7 µm (pore size) Whatman GF/F glass fiber filters (pre-combusted at 400 °C for 4 h in a Muffle furnace) and then through 0.22 µm (pore size) nuclepore filters (Whatman) [2]. Then spectrophotometer (Shimadzu UV-2600) was used to measure the CDOM absorbance spectra (OD($\lambda$)) between 200 and 800 nm at 1 nm intervals with the filtrate in 1 cm quartz cuvette and Milli-Q water as reference. The absorption spectrum ($a_{\text{CDOM}}(\lambda)$) was calculated from the absorbance using Equation (3) [46], where L is the cuvette path length (0.01 m) and 2.303 is the conversion factor. Some fine particles may
have remained in the filtrate so backscattering caused by them should be corrected with absorption of CDOM was assumed to be zero at $\lambda_0$, where $\lambda_0$ equals to 700 nm [2].

$$a'_{\text{CDOM}}(\lambda) = \frac{2.303}{L} \times [\text{OD}(\lambda) - \text{OD}(\lambda_0) \times (\lambda/\lambda_0)]$$  \hfill (3)

The absorption coefficient at 355 nm ($a_{\text{CDOM}(355)}$) was selected to represent the CDOM concentration [47,48].

Absorption coefficients of phytoplankton ($a_{\text{ph}}$) and NAP ($a_{\text{NAP}}$) were measured by the quantitative filter technique (QFT). A certain volume of each water sample was filtered through Whatman GF/F filter with a nominal pore size of 0.7 $\mu$m. Absorption of total particles ($a_p$) on the filter was measured by spectrophotometer (Shimadzu UV-2600), and then $a_{\text{NAP}}$ was measured after the filter was bleached by sodium hypochlorite solution to remove the pigment. As for phytoplankton absorption, $a_{\text{ph}}$ was calculated by subtracting $a_{\text{NAP}}$ from $a_p$. The details can be found in [40,49]. Due to the artificial factors during the experiment, only $a_p$ of each sample point were measured for Baishan Reservoir (BSR number 4, Table 1).

| Table 1. Summaries of sample points and average values of optical parameters of the sampling lakes. N: counts of sample points; SDD: Secchi disk depth; TSM: total suspended matter; $a_{\text{CDOM}}(355)$: the absorption coefficient of chromophoric dissolved organic matter (CDOM) at 355 nm; Chl-$a$: chlorophyll-$a$ concentration. |
|-----------------|----------------|---------|--------|--------|--------|--------|--------|--------|
| Water Name      | Abbreviation   | Area (km$^2$) | Date    | N      | $K_p$(PAR m$^{-1}$) | SDD (m) | TSM (mg/L) | $a_{\text{CDOM}}(355)$ m$^{-1}$ | Chl-$a$ (µg/L) |
| Shanmen R.      | SMR(1)         | 1.5      | 21 April 2015 | 4     | 0.77                | 1.543    | 3.17    | 7.86               | 5.88               |
| Xiaotantai R.   | XST(2)         | 1.3      | 21 April 2015 | 4     | 2.73                | 0.703    | 18.75   | 10.56              | 32.84              |
| Xinmo Island    | XM(3)          | 26.6     | 24 April 2015 | 8     | 2.53                | 0.506    | 26.01   | 3.51               | 7.66               |
| Baishan R.      | BSR(4)         | 90.0     | 4 May 2015    | 24    | 1.11                | 1.283    | 6.29    | 5.08               | 28.02              |
| Xiaomeng R.     | XMK(5)         | 162.1    | 15 August 2015 | 13    | 3.65                | 0.362    | 34.77   | 4.43               | 6.21               |
| Dazhang R.      | DZK(6)         | 1062.4   | 17 August 2015 | 8     | 5.43                | 0.226    | 55.17   | 1.85               | 6.58               |
| Qingyin R.      | QYN(7)         | 41.1     | 18 August 2015 | 13    | 13.93               | 0.113    | 174.50  | 5.14               | 4.31               |
| Lianhua R.      | LHL(8)         | 111.7    | 19 August 2015 | 14    | 1.80                | 1.284    | 46.1   | 4.84               | 23.00              |
| Jingbo L.       | JB(9)          | 88.8     | 20 August 2015 | 11    | 1.22                | 1.520    | 4.18    | 2.44               | 5.41               |
| Songhua L.      | SSL(10)        | 216.2    | 21 August 2015 | 7     | 0.68                | 2.599    | 8.12    | 4.30               | 16.89              |
| Kulipao         | KL(11)         | 11.6     | 6 September 2015 | 4     | 4.32                | 0.328    | 22.10   | 10.36              | 3.74               |
| Nanyin R.       | NYR(12)        | 96.4     | 7 September 2015 | 3     | 5.35                | 0.253    | 34.17   | 3.22               | 40.04              |
| Lamioping       | LM(13)         | 47.9     | 8 September 2015 | 7     | 2.81                | 0.440    | 16.10   | 6.03               | 59.80              |
| Longhuipo       | LH(14)         | 126.9    | 9 September 2015 | 10    | 4.11                | 0.280    | 36.33   | 4.69               | 12.27              |
| Talahengpo      | TL(15)         | 67.5     | 10 September 2015 | 2    | 5.63                | 0.240    | 56.00   | 6.49               | 20.62              |
| Xihulipo        | XH(16)         | 57.9     | 10 September 2015 | 7    | 6.23                | 0.254    | 50.00   | 4.71               | 21.25              |
| Huashanpo       | HS(17)         | 64.3     | 10 September 2015 | 8     | 4.00                | 0.371    | 27.18   | 4.53               | 9.02               |
| Hulan L.        | HL(18)         | 2050.2   | 14 September 2015 | 28    | 3.54                | 0.402    | 28.38   | 5.23               | 7.37               |
| Nierui R.       | NEIR(19)       | 429.6    | 16 September 2015 | 16    | 1.56                | 1.678    | 4.24    | 6.53               | 4.46               |
| Shandu R.       | SKR(20)        | 64.9     | 17 September 2015 | 3     | 1.27                | 1.847    | 1.99    | 7.59               | 6.88               |
| Nancangzi R.    | NZR(21)        | 10.7     | 14 April 2016   | 4     | 0.85                | 1.80     | 1.98    | 2.45               | 2.97               |
| Qinghe R.       | QR(22)         | 17.3     | 15 April 2016   | 3     | 1.49                | 1.24     | 8.61    | 1.98               | 5.03               |
| Chahe R.        | CHR(23)        | 12.1     | 15 April 2016   | 3     | 1.00                | 1.17     | 4.00    | 1.83               | 8.12               |
| Tanghe R.       | THR(24)        | 18.1     | 18 April 2016   | 4     | 0.92                | 1.85     | 4.47    | 1.07               | 3.25               |
| Huashen R.      | HR(25)         | 69.6     | 19 April 2016   | 2     | 0.47                | 2.74     | 2.63    | 1.92               | 3.58               |
| Shuifeng R.     | SF(26)         | 165.7    | 20 April 2016   | 5     | 0.97                | 1.45     | 8.85    | 1.98               | 3.94               |

$a$ denotes that L. = Lake; R. = Reservoir; pao means lakes in Northeast China; LHL, JBL and SHL are actually reservoirs but their Chinese names are called lakes.

2.4. Grey Incidences (GIs) Analysis

The grey system theory is a method of processing and analyzing systems with incomplete information [50]. It has been used in remote sensing applications [51]. GIs provide a quantitative description of the system and dominant factor which influences the system’s development can be determined with GIs. Given a system contains m factors and one output, for k times of tests of the system, it generates a data sequence with k values of the output as Equation (4). This data sequence

$$G = \left[ \begin{array}{ccc} k_1 & k_2 & \cdots \cdots & k_m \\ x_1 & x_2 & \cdots \cdot & x_k \\ \vdots & \vdots & \ddots & \vdots \\ x_1 & x_2 & \cdots \cdot & x_k \end{array} \right]$$  \hfill (4)
is defined as systematic behavior. The corresponding factors compose m data sequences that each contains k values (Equation (5)). This data sequences are defined as factor’s behavior.

\[ X_0 = \{x_0(k), k = 1, 2 \ldots n\} \] (4)

\[ X_i = \{x_0(i), k = 1, 2 \ldots n\} \quad (i = 1, 2, 3 \ldots m) \] (5)

In Equations (4) and (5), \( X_0 \) and \( X_i \) represent the data sequences, \( x_0(k) \) is the system output at kth test, and \( x_i(k) \) is system factor value of ith factor at kth test. GIs describe the developmental trends of behaviors and factors in a system by analyzing the similarity of geometric patterns between systematic behavior and factor’s behavior data sequences [50]. By calculating and comparing GIs between \( X_0 \) and \( X_i \), the most influential factors could be identified.

In this research, \( K_d(PAR) \) of the sample points was regarded as the systematic behavior meanwhile corresponding Chl-\( a \), TSM, and CDOM were regarded as factor’s behavior. Three types of GI were calculated. The first one (GI1) was proposed by Deng [52] and the details about calculation can be found in [53]. GI1 was the first model proposed in grey system theory. It measured the trend of system behaviors and factors by the distance between corresponding points of the data sequences. The second one (GI2) was the improved generalized absolute grey incidence model proposed by Cao and the details about its calculation can be found in [50]. GI2 is the modification of the GI proposed by Liu [53]. It measured the trend of system behaviors and factors by the area of the region that surrounded by the curves of the data sequences. Finally, the third one (GI3) was the absolute degree of grey incidence proposed by Mei [54]. According to GI3 model, data sequences \( X_0 \) and \( X_i \) were converted to \( Y_0 = \{y_0(k), k = 1, 2, \ldots ,n - 1\} \) and \( Y_i = \{y_i(k), k = 1, 2, \ldots ,n - 1\} \), respectively, using Equation (6). This conversion calculates the slopes of the adjacent data points in each data sequence.

\[ y(k) = x(k + 1) - x(k) \quad k = 1, 2, 3, \ldots ,n - 1 \] (6)

\[ \text{GI}(X_0, X_i) = \frac{1}{n - 1} \sum_{i=1}^{n-1} \frac{1}{1 + |y_0(k) - y_i(k)|} \] (7)

The GI of two data sequences (GI\( (X_0,X_i) \)) was calculated from Equation (7). As shown above, the GI3 measured the trend of system behaviors and factors by the slope of the data sequences.

In order to eliminate influence of dimension, the data sequences were standardized to values ranging from 0 to 1 according to Equation (8) before calculating the three kinds of GIs.

\[ x_i'(k) = \frac{x_i(k) - \min(X)}{\max(X) - \min(X)} \] (8)

2.5. Linear Regression between \( K_d(PAR) \) and OACs

GIs provided the means to identify which of the OACs acts as the dominant factor in determining \( K_d(PAR) \), however it does not provide a relationship to quantitatively derive \( K_d(PAR) \) from the factor. Therefore, a linear regression analysis was performed to establish the quantitative relationship between \( K_d(PAR) \) and the concentration of the OACs. In some researches the regression analysis was also used to identify the dominant factor of \( K_d(PAR) \) [2,7] so it can also be used to validate the result of GIs analysis in this paper. In order to further evaluate the applicability of the predicting model, we first used the leave-one-out cross validation (LOO-CV) method and then accessed the model’s spatial transferability. For the LOO-CV analysis, \( n - 1 \) samples were used to calibrate \( K_d(PAR) \) and the sample left out was used to validate the model. Similarly, leave one lake out cross validation was used for the assessment of spatial transferability, sample points from \( n - 1 \) lakes were used to calibrate the \( K_d(PAR) \) and the sample points of the lake that was left out were used for validation. The prediction error sum of squares (PRESS) was used to derive the root-mean-square error of cross-validation (RMSECV) of the LOO-CV. Moreover, the relative error (RE), the mean relative error (MRE), and the root-mean-square
error (RMSE) were used to assess the accuracy of the model. The details on how to calculate these indicators can be found in [12,44].

2.6. Relationship between $K_d$(PAR) and SDD

Water transparency is an easily measured indicator of water quality and is used to calculate the trophic state index [55]. SDD is correlated with $K_d$(PAR) as they are both influenced by the absorption and scatter characteristics of the OACs; in effect, both measurements represent the penetration of light in the water [56]. The difference between them is the varied contributions of the extent of spectral bands as SDD is related to the visible domain (410–665 nm) [57,58] and $K_d$(PAR) is 400–700 nm [5,33]. $K_d$(PAR) and SDD are inversely correlated since higher $K_d$(PAR) values indicate lower water clarity. The relationship of SDD and $K_d$(PAR) as defined by Holmes [8] is:

$$K_d(PAR) = f \times SDD^{-1}$$

where $f$ has the value of 1.44 for turbid coastal waters [8]. However, other studies have reported different values of $f$, ranging from 1.7 to 2.3 [59,60]. In this paper, $f$ was determined by linear regression with a fixed intercept at 0. With the relationship $K_d$(PAR) can be estimated from SDD.

3. Results and Discussion

3.1. Water Quality Characteristics

Due to the different geographical environments, the sampled area included a large diversity of inland waters with varying concentration of OACs (Table 1). The concentration of Chl-$a$ (average: $15.16 \pm 15.78 \mu g/L$) varied from 1.20 (sample point number 10 of NEJR, number 19, Table 1) to $67.15 \mu g/L$ (sample point number 1 of LMSP, number 13, Table 1). BSR (number 4, Table 1) exhibited the largest variation of Chl-$a$ ranging from 6.36 to 60.94 $\mu g/L$ (average: $28.02 \pm 20.35 \mu g/L$) attributed to its long-narrow shape. Upstream regions contained high concentration of Chl-$a$. As the water depth becomes deep and flow velocity becomes slow from upstream to downstream regions, the suspended components sunken so concentration of Chl-$a$ gradually decreased in the up layer of the water. The TSM concentration ranged from 0.83 (sample point number 3 of SHL, number 10, Table 1) to 184 mg/L (sample point number 4 of QNR, number 7, Table 1), with an average of $4.55 \pm 28.25 \mu g/L$. The highest concentration appeared in QNR (average: $174.5 \pm 8.85 \mu g/L$) and the lowest was in SHL (average: $1.48 \pm 0.97 \mu g/L$). XMP (number 3, Table 1) had the largest variation in the concentration of TSM (range: 10.75 to 64 mg/L, average: $26.01 \pm 17.32 \mu g/L$). This variation resulted from one particular sample point with very high TSM concentration (64 mg/L). It was collected in a site with high water turbulence as it located at the junction of a river and the lake. The absorption coefficient of CDOM at 355 nm was high both in XSTR ($10.56 \pm 0.25 m^{-1}$, number 2, Table 1) and KLP ($10.56 \pm 0.44 m^{-1}$, number 11, Table 1) due to their grayish yellow water color.

The SDD ranged from 0.1 to 4.32 m (Table 1). QNR (number 7, Table 1) had the lowest water clarity ($0.11 \pm 0.01 m$), while HRR (number 25, Table 1) was the clearest ($2.74 \pm 0.34 m$). The coefficient $f$ fitted from all sample points was equal to 1.38 ($R^2 = 0.97$) (Figure 2a). This result was closer to the one reported by Holmes [8] indicating that the equation proposed therein can be used to predict the $K_d$(PAR) from the SDD measurements.

3.2. Spatial Distribution of $K_d$(PAR)

The mean $K_d$(PAR) of each lake and their corresponding geographical distribution are shown in Figure 1b. In general, the reservoirs displayed low $K_d$(PAR) values. The reason might be due to the fact that they are located in mountainous areas, which are covered by forests that prevent soil erosion, hence, resulting in low concentrations of TSM. Lakes distributed in the Songnen Plain overall displayed high $K_d$(PAR) values. The soil erosion around these lakes was strong, the lakes were shallow,
and the re-suspension caused by strong winds in the area also contributed to high turbidity [39]. These findings were in agreement with the research conducted by Olmanson [61], which revealed that lakes of Minnesota that were located in forest regions were clearer than those located in plain regions.

Figure 2. Correlation between $K_d$(PAR) and the concentration of the water’s quality parameters: (a) water transparency (SDD); (b) total suspended matter (TSM); (c) chlorophyll-a (Chl-a); and (d) absorption coefficient of chromophoric dissolved organic matter (CDOM) at 355 nm ($a_{CDOM}(355)$).

3.3. Grey Incidences between $K_d$(PAR) and OACs

According to previous studies [48], absorption coefficient at 355 nm of CDOM ($a_{CDOM}(355)$) was regarded as a representative measure of the concentration of CDOM. GIs were calculated with TSM, Chl-a and $a_{CDOM}(355)$ as the factor’s behavior data sequences while $K_d$(PAR) as the systematic behavior data sequence. The factor with higher GIs score indicates a relative bigger influence on $K_d$(PAR). The GIs of all the sample points were calculated and the results (Table 2) showed that the highest GIs were obtained by comparing between TSM and $K_d$(PAR). These results indicated that TSM had a higher contribution to $K_d$(PAR) than Chl-a and CDOM, which is consistent with previous studies in Lake Taihu in China [7,12] and UK marine waters [10]. Furthermore, in order to assess the difference of the relationships across various waters, lakes with more than 10 sampling points were also selected to calculate GIs individually. The results (Table 2) indicated that the dominant factors of $K_d$(PAR) varied according to lakes. For example, $K_d$(PAR) of BSR may largely influenced by TSM and Chl-a as the GIs of two materials did not vary significantly.
Table 2. GIs between the OACs and $K_d$(PAR) of all the sample points and of lakes whose number of sampling points were greater than 10. The bigger the GI the larger the influence of OACs on $K_d$(PAR).

<table>
<thead>
<tr>
<th>Name</th>
<th>OACs</th>
<th>GI 1</th>
<th>GI 2</th>
<th>GI 3</th>
<th>Name</th>
<th>OACs</th>
<th>GI 1</th>
<th>GI 2</th>
<th>GI 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>All points</td>
<td>TSM</td>
<td>0.82</td>
<td>0.81</td>
<td>0.97</td>
<td>Jingbo L.</td>
<td>TSM</td>
<td>0.73</td>
<td>0.73</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>$a_{CDOM}(355)$</td>
<td>0.51</td>
<td>0.69</td>
<td>0.86</td>
<td></td>
<td>$a_{CDOM}(355)$</td>
<td>0.63</td>
<td>0.87</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>Chl-$a$</td>
<td>0.63</td>
<td>0.66</td>
<td>0.85</td>
<td></td>
<td>Chl-$a$</td>
<td>0.58</td>
<td>0.81</td>
<td>0.83</td>
</tr>
<tr>
<td>Baishan R.</td>
<td>TSM</td>
<td>0.77</td>
<td>0.91</td>
<td>0.92</td>
<td>Longhupao</td>
<td>TSM</td>
<td>0.73</td>
<td>0.65</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>$a_{CDOM}(355)$</td>
<td>0.47</td>
<td>0.78</td>
<td>0.82</td>
<td></td>
<td>$a_{CDOM}(355)$</td>
<td>0.66</td>
<td>0.45</td>
<td>0.63</td>
</tr>
<tr>
<td></td>
<td>Chl-$a$</td>
<td>0.73</td>
<td>0.87</td>
<td>0.91</td>
<td></td>
<td>Chl-$a$</td>
<td>0.71</td>
<td>0.53</td>
<td>0.82</td>
</tr>
<tr>
<td>Xiaoxingkai L.</td>
<td>TSM</td>
<td>0.83</td>
<td>0.77</td>
<td>0.88</td>
<td>Hulan L.</td>
<td>TSM</td>
<td>0.73</td>
<td>0.54</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>$a_{CDOM}(355)$</td>
<td>0.57</td>
<td>0.47</td>
<td>0.76</td>
<td></td>
<td>$a_{CDOM}(355)$</td>
<td>0.68</td>
<td>0.45</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>Chl-$a$</td>
<td>0.73</td>
<td>0.74</td>
<td>0.74</td>
<td></td>
<td>Chl-$a$</td>
<td>0.65</td>
<td>0.30</td>
<td>0.75</td>
</tr>
<tr>
<td>Lianhua L.</td>
<td>TSM</td>
<td>0.86</td>
<td>0.93</td>
<td>0.92</td>
<td>Nierji R.</td>
<td>TSM</td>
<td>0.72</td>
<td>0.71</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td>$a_{CDOM}(355)$</td>
<td>0.55</td>
<td>0.71</td>
<td>0.89</td>
<td></td>
<td>$a_{CDOM}(355)$</td>
<td>0.62</td>
<td>0.56</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>Chl-$a$</td>
<td>0.66</td>
<td>0.72</td>
<td>0.77</td>
<td></td>
<td>Chl-$a$</td>
<td>0.57</td>
<td>0.46</td>
<td>0.87</td>
</tr>
</tbody>
</table>

L = Lake; R = Reservoir.

3.4. $K_d$(PAR) Model Calibration and Validation

In this study, TSM was revealed to be the major impact factor for the determination of $K_d$(PAR). The linear regression analysis provided the means to predict $K_d$(PAR) by utilizing the OACs data. The results indicated that $K_d$(PAR) was strongly correlated to TSM, while there was no obvious relationship between $K_d$(PAR) and Chl-$a$ or CDOM (Figure 2). The mean value of $K_d$(PAR) and OACs of each lake were determined by averaging all the sample points collected from the corresponding lakes. The coefficient of determination ($R^2$) was improved when averaged value of each lake, rather than all the single point values, was fitted in the linear model. This is because experimental error may exist in a single sample point, which leads to the dispersion of the data. The average of the data can eliminate the error to some extent and result in a stable performance. The slope of the linear model between $K_d$(PAR) and TSM in the current study was slightly higher than results in [7] (slope = 0.0626, intercept = 1.6068) and [12] (slope = 0.0563, intercept = 1.52). Considered that our research covered a larger number of lakes than those earlier studies, such a difference was acceptable. The results of linear regression also indicated that the dominant factor of $K_d$(PAR) for lakes of Northeastern China is TSM as the GIs indicated. This result provides a guidance in band selection to derive $K_d$(PAR) from remote sensing image that the bands well correlate to TSM may perform well in deriving $K_d$(PAR) [12]. The model estimating $K_d$(PAR) from TSM may be applied in mapping $K_d$(PAR) indirectly from remote sensing image. The dominant factor of $K_d$(PAR) for lakes in Northeastern China was same to Lake Taihu, a large shallow lake in Eastern China [2,7,12]. However, the coefficients of the linear relationship indicated there was difference in optical properties between Lake Taihu and lakes in Northeastern China.

A multivariate linear regression analysis was also performed with $K_d$(PAR) as dependent variable and TSM, Chl-$a$ and $a_{CDOM}(355)$ as independent variables. A slight better fit was obtained as the $R^2$ was 0.916 for all sample point values and 0.960 for lake specified mean values. The result was similar to those of Zhang [2] and Devlin [11], demonstrating that slight improved models were achieved when fit $K_d$(PAR) with all three explanatory variables rather than solely TSM for waters that $K_d$(PAR) was mainly influenced by TSM.

The simple linear model may be suitable for predicting $K_d$(PAR) from TSM as it produced low error rates (RMSE: 0.689 for lake averaged values, and 0.709 for single point values). The RMSECV of the LOO-CV analysis was 0.709 and the MAPE was 0.315. The RE which was used for validation ranged from 0.0016 to 2.2 (Figure 3). The relative error of 61% of the samples was below 0.3 and of 74% of the samples was below 0.4. Altogether our results indicated that the TSM performed well in estimating $K_d$(PAR). However, care should be taken when applying this model to some cases as TSM might not always represent the major determining factor of $K_d$(PAR). This could be the case of eight sample points in our study which displayed RE values larger than 1.0 in LOO-CV.
was dominated by CDOM and phytoplankton which might cause the large error observed by solely estimating $K_d$(PAR) from the concentration of TSM. However, when compared to the model built by all sample points, the slope and intercept decreased to 0.06 and 0.522, respectively. This explained why the model built with all sample points lead to overestimation of $K_d$(PAR) for the 32 sample points. However, the GIs revealed that the correlation between the OACs (TSM, Chl-$a$, CDOM) and $K_d$(PAR) did not vary significantly (GI1: 0.70, 0.65, 0.68), (GI2: 0.55, 0.67, 0.72) and (GI3: 0.91, 0.80, 0.80), which indicated that the dominant factors of $K_d$(PAR) may not solely TSM for this 32 sample points. The $R^2$ of multivariate linear regression analysis was 0.64, which implied an improved fit with TSM, Chl-$a$ and CDOM.

To investigate this further, 32 points, whose RE of LOO-CV was greater than 0.6, were selected to analyze the relationship between $K_d$(PAR) and OACs. For the 32 sample points, the averaged $K_d$(PAR) was $0.727 \pm 0.150 \text{ m}^{-1}$ and the large RE value was mainly caused by the overestimation of the model (see Figure S2). The linear regression analysis (see Figure S3) indicated that $K_d$(PAR) had a better correlation with the concentration of TSM ($R^2 = 0.60$) than with Chl-$a$ and CDOM ($R^2 < 0.1$). However, the integration of the absorption coefficient curves between 400–700 nm were used to calculate the relative contribution of phytoplankton, NAP and CDOM to the total absorption of the samples (Figure 4). Due to the absorption spectra of phytoplankton and NAP in the BSR were not measured, therefore, six data points from BSR were not analyzed further. Nonetheless, the absorption spectra of the total particulate matter showed an obvious peak at 675 nm (see Figure S4), which indicated that the contribution of the phytoplankton to the total absorption is strong. The total absorption of HRR was dominated by CDOM and phytoplankton which might cause the large error observed by solely estimating $K_d$(PAR) from TSM. Though NAP contributed more than phytoplankton and CDOM to the total absorption of QHR1, CHR1, THR and SFR, high RE were observed when predicting $K_d$(PAR) solely from the TSM. This might be attributed to the extremely clear water which makes the model unsuitable for this type of environment; therefore, a separate analysis should be done for this lake. For JBL and SHL, the CDOM was the main contributor to the total absorption, thus, the $K_d$(PAR) may have a good correlation with it, as discussed in Section 3.5.

Figure 3. Frequency distribution and cumulative percentage of the relative errors (RE) derived from the leave-out-one cross validation (LOO-CV) of the regression model that had been used to calibrate the $K_d$(PAR) from the concentration of TSM.
were low for HHR (LakeID: 25), SFR (LakeID: 26), SHL (LakeID: 10) SMR (LakeID: 1), THR (LakeID: 24), JBL (LakeID: 9) and BSR (LakeID: 4) (Figure 5), the MRE (MRE: 1.794, 0.862, 0.860, 0.726, 0.694, 0.641 and 0.532, respectively) indicated that there were large errors in predicting \( K_d(PAR) \) in this lakes. These lakes commonly had low levels of TSM, so the relative contributions of Chl-\( \alpha \) and CDOM could not be eliminated from the analysis as they may also have large influence on \( K_d(PAR) \). Otherwise, it may cause the high MRE values in predicting \( K_d(PAR) \) solely with TSM.

In order to identify the differential contribution of OACs to the estimation of \( K_d(PAR) \) across different lakes, a separate analysis was undertaken for lakes with more than 10 sample points. Although only six sample points were obtained from SHL, it was still selected because it was the second clearest lake in this study, and it had RE greater than 0.4 in the LOO-CV analysis. The GIs for the lakes were calculated respectively and linear regression between OACs and \( K_d(PAR) \) was performed. The lake specified averaged absorption spectra were used to calculate the relative contribution of phytoplankton, NAP and CDOM to the total absorption of each lake. The \( R^2 \) of linear regression and the relative contribution were shown in Figure 6.

GIs of the lakes are shown in Table 2. For SHL, the GIs between the OACs (TSM, Chl-\( \alpha \), CDOM) and the \( K_d(PAR) \) are (GI1: 0.62, 0.81, 0.67), (GI2: 0.77, 0.90, 0.80) and (GI3: 0.77, 0.89, 0.83). The GIs of SHL showed that Chl-\( \alpha \) and CDOM displayed a higher correlation than TSM with the \( K_d(PAR) \), and this was consistent with the results obtained by linear regression analysis. The contribution of CDOM and phytoplankton to the total absorption was larger than NAP and this could explain why \( K_d(PAR) \) was mainly influenced by Chl-\( \alpha \) and CDOM. The \( R^2 \) of the correlation analysis revealed that the \( K_d(PAR) \) of BSR was highly related to both TSM and Chl-\( \alpha \), and that there was only a minor variation in the GIs of TSM and Chl-\( \alpha \). The \( K_d(PAR) \) of NEJR was greatly correlated to both TSM.

3.5. Spatial Transferability

The MRE and RMSE of each lake were calculated in the spatial transferability assessment (Figure 5). QNR (LakeID: 7), KLP (LakeID: 11), NYR (LakeID: 12), and XHLP (LakeID: 16) all had high RMSE values but relatively low MRE values due to their high \( K_d(PAR) \). Unlike the RMSE values, which were low for HHR (LakeID: 25), SFR (LakeID: 26), SHL (LakeID: 10) SMR (LakeID: 1), THR (LakeID: 24), JBL (LakeID: 9) and BSR (LakeID: 4) (Figure 5), the MRE (MRE: 1.794, 0.862, 0.860, 0.726, 0.694, 0.641 and 0.532, respectively) indicated that there were large errors in predicting \( K_d(PAR) \) in this lakes. These lakes commonly had low levels of TSM, so the relative contributions of Chl-\( \alpha \) and CDOM could not be eliminated from the analysis as they may also have large influence on \( K_d(PAR) \). Otherwise, it may cause the high MRE values in predicting \( K_d(PAR) \) solely with TSM.

In order to identify the differential contribution of OACs to the estimation of \( K_d(PAR) \) across different lakes, a separate analysis was undertaken for lakes with more than 10 sample points. Although only six sample points were obtained from SHL, it was still selected because it was the second clearest lake in this study, and it had RE greater than 0.4 in the LOO-CV analysis. The GIs for the lakes were calculated respectively and linear regression between OACs and \( K_d(PAR) \) was performed. The lake specified averaged absorption spectra were used to calculate the relative contribution of phytoplankton, NAP and CDOM to the total absorption of each lake. The \( R^2 \) of linear regression and the relative contribution were shown in Figure 6.

GIs of the lakes are shown in Table 2. For SHL, the GIs between the OACs (TSM, Chl-\( \alpha \), CDOM) and the \( K_d(PAR) \) are (GI1: 0.62, 0.81, 0.67), (GI2: 0.77, 0.90, 0.80) and (GI3: 0.77, 0.89, 0.83). The GIs of SHL showed that Chl-\( \alpha \) and CDOM displayed a higher correlation than TSM with the \( K_d(PAR) \), and this was consistent with the results obtained by linear regression analysis. The contribution of CDOM and phytoplankton to the total absorption was larger than NAP and this could explain why \( K_d(PAR) \) was mainly influenced by Chl-\( \alpha \) and CDOM. The \( R^2 \) of the correlation analysis revealed that the \( K_d(PAR) \) of BSR was highly related to both TSM and Chl-\( \alpha \), and that there was only a minor variation in the GIs of TSM and Chl-\( \alpha \). The \( K_d(PAR) \) of NEJR was greatly correlated to both TSM.

Figure 4. Relative contribution of phytoplankton (\( a_{ph} \)), NAP (\( a_{NAP} \)) and CDOM (\( a_{CDOM} \)) to the total absorption of the sample points with relative errors (RE) greater than 0.6 (based on LOO-CV analysis). Sample points are represented on the x-axis with the abbreviations of the location’s name.
and CDOM, and CDOM and NAP contributed more to the total absorption. This was similar to the analyzing results of XXKH. Overall, the count of dominant factors for determining $K_d$(PAR) were two rather one in the cases of the lakes/reservoirs described above. This results is consistent with the findings in Danjiangkou Reservoir (averaged $K_d$(PAR) was 0.726 m$^{-1}$) in Hubei Province, China [32]. That study revealed that TSM and Chl-$a$ dominated $K_d$(PAR) during the wet season while Chl-$a$ and CDOM were the major determinants of $K_d$(PAR) in the dry season in the reservoir.

Figure 5. The mean relative error (MRE) and root-mean-square error (RMSE) of the $K_d$(PAR) regression models for each lake for the evaluation of the spatial transferability of the model. LakeID was corresponding to Table 1 and was represented on the x-axis.

Figure 6. The bars represent the determination coefficients ($R^2$) of the linear regression analysis between the OACs and the $K_d$(PAR). The lines represent the relative contribution of phytoplankton ($\alpha_{ph}$), non-algal particles ($\alpha_{NAP}$) and CDOM ($\alpha_{CDOM}$) to the total absorption. Abbreviations of the lakes correspond to those in Table 1 and are represented on the x-axis.

The TSM and CDOM concentrations were well correlated to $K_d$(PAR) for JBL. Similarly, Chl-$a$ also had a modest correlation with $K_d$(PAR). In contrast to JBL which displayed similar contributions of each constituent to the total absorption, the $K_d$(PAR) of LHL was mainly correlated to the TSM.
The reason may be that the TSM was more varied in LHL (range: 2.83–22.6 mg/L, StDev: 6.33) than in JBL (range: 2.14–10.2 mg/L, StDev: 2.3), while the Chl-a was more varied in JBL (range: 9.94–61.67 μg/L, StDev: 14.61) than in LHL (range: 6.81–43.11 μg/L, StDev: 10.43). Though NAP contributed the most to the total absorption in LHP and HLH, \( K_d \) (PAR) had a weak correlation with the TSM. Taking into account that these two lakes are vast in shape and relative shallow in depth (5.92 m for HLH and 2.7 m for LHP) [37], it is possible that the OACs are uniformly distributed in the water causing a little variation in the calculation of \( K_d \) (PAR) across different sample points. This might have produced low standard deviations in the values of the dependent and the independent variables, in effect, lowering the significance of the linear correlation. In addition, the measurement of PAR was easily influenced by waves in these lakes which could induce large errors to the calculation of \( K_d \). The optical properties of water bodies were diverse and complex. \( K_d \) was easily influenced by waves in these lakes which could induce large errors to the calculation of \( K_d \). As a systemic analysis method, GIs were effective in identifying the dominant factors of \( K_d \) (PAR). (5.92 m for HLH and 2.7 m for LHP) it is possible that the OACs are uniformly distributed in the water causing a little variation in the calculation of \( K_d \) (PAR). This inconformity was also observed by Shi [29] in Bosten Lake in Xinjiang Province, China. In his study, both CDOM and phytoplankton contributed more than NAP to the total absorption, however, the \( K_d \) (PAR) was best correlated to the TSM concentration.

4. Conclusions

Based on the data collected in Northeast China, the relationships between \( K_d \) (PAR) and OACs were studied. The optical properties of water bodies were diverse and complex. \( K_d \) (PAR) was significantly correlated to the water transparency (SDD) and the coefficient was in agreement with previous studies [8]. According to GIs and linear regression analysis of the data of all the sampled points, it was found that TSM was the most dominant component of the OACs in regulating \( K_d \) (PAR). As a systemic analysis method, GIs were effective in identifying the dominant factors of \( K_d \) (PAR). The TSM accounted for the most variation of the \( K_d \) (PAR) as the \( R^2 \) of the linear model only increased 0.01 when TSM, Chl-a, and CDOM were the explanatory variables rather than solely TSM was the explanatory variable. The investigation of the simple linear model demonstrated that it was an effective tool to predict \( K_d \) (PAR) from TSM data as it produced small RE in a LOO-CV analysis. However, the dominant regulating factors of \( K_d \) (PAR) varied among some of the sample points and lakes, especially for some clear waters, in which the dominant factors may be a combination of two or three kinds of OACs. This variation can be attributed to their distinct geographical environments. Lakes that are open in shape and shallow in depth and distributed in plain areas, where soil erosion is relatively strong, show high \( K_d \) (PAR) values. Therefore, the major determining factor of \( K_d \) (PAR) in these types of lakes is the concentration of TSM. However, for very clear water bodies, for example in deep and narrow reservoirs, which are distributed in mountainous areas where soil erosion is weak, the concentration of the TSM is low. In such cases, the contribution of the concentrations of Chl-a and/or CDOM to the estimation of \( K_d \) (PAR) is also significant. There was a discrepancy in the results derived from correlation analysis and from absorption spectra analysis. The material which contributed the most to the total absorption may not be well correlated to the \( K_d \) (PAR) due to the scattering of phytoplankton and TSM also affect \( K_d \) (PAR) [29]. In general, reservoirs in the east part of Northeast China had low \( K_d \) (PAR) values, while lakes located in plain areas showed high \( K_d \) (PAR) values. These differences in the \( K_d \) (PAR) could reflect variations in the composition of TSM in these distinct geographical locations.

Supplementary Materials: The following are available online at www.mdpi.com/2072-4292/8/11/964/s1, Figure S1: Distribution of sample points of each lakes or reservoirs. The color represents the average \( K_d \) (PAR), the number and abbreviation were listed in Table 1; Figure S2: The estimated \( K_d \) (PAR) by TSM and in situ \( K_d \) (PAR) for sample points that had relative error (RE) greater than 0.6 in LOO-CV. Eight sample points had RE greater than 1 and they were HRR1 (RE: 2.20), JBH1 (RE: 1.59), JBH3(RE: 1.54), SHH7(RE: 1.50), SFR5(RE: 1.49), HRR2(RE: 1.36), SHH5(RE: 1.15) and QHR1(RE: 1.06); Figure S3: Linear regression analysis between OACs and \( K_d \) (PAR) for sample points that had relative error (RE) greater than 0.6 in LOO-CV. (a) total suspended matter (TSM); (b) chlorophyll-a(Chl-a); (c) absorption coefficient of chromophoric dissolved organic matter (CDOM) at 355 nm (aCDOM(355)); Figure S4: Absorption coefficient of total particulate materials of BSR. Only the absorption coefficients of total particulate materials were measured during the experiment but the phytoplankton and...
non-algal particles were not measured. The obvious absorption peak at 675 nm indicated the pigment absorption was dominant in the total particulate absorption.

Acknowledgments: The research was jointly supported by the “One Hundred Talents” program from Chinese Academy of Sciences granted to Kaishan Song, the National Natural Science Foundation of China (No. 41471290, No. 41501387) and the Foundation of China’s High Resolution Earth Observation (No. 41-Y20A31-9003-15/17). The authors thank Jia Du, Ming Wang, and Junbin Hou for their persistent assistance with field sampling and in situ measurements. The authors also thank the editors and three anonymous reviewers for their valuable and instructive comments that have strengthened the manuscript greatly.

Author Contributions: Kaishan Song conceived the framework and designed the fieldwork; Zhidan Wen, Ying Zhao, Yingxin Shang, Chong Fang, Jia Du, and Jianhang Ma performed the fieldwork and experiments; Jianhang Ma analyzed the data and wrote the paper; and Jianhang Ma and Kaishan Song revised the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References


56. Aas, E.; Høkedal, J.; Sørensen, K. Secchi depth in the Oslofjord-Skagerrak area: Theory, experiments and relationships to other quantities. *Ocean Sci.* 2014, 10, 177–199. [CrossRef]


