Phytoplankton size class in the East China Sea derived from MODIS satellite data

Hailong Zhang[1,2], Shengqiang Wang[1,2], Zhongfeng Qiu[1,2], Deyong Sun[1,2], Joji Ishizaka[3], Shaojie Sun[4], Yijun He[1,2]

[1] School of Marine Sciences, Nanjing University of Information Science & Technology, Nanjing, Jiangsu, China

[2] Jiangsu Research Centre for Ocean Survey Technology, NUIST, Nanjing, Jiangsu, China

[3] Institute for Space-Earth Environmental Research, Nagoya University, Nagoya, Japan

[4] College of Marine Science, University of South Florida, St. Petersburg, Florida, USA

Correspondence to: Z. F. Qiu (zhongfeng.qiu@nuist.edu.cn)

Abstract. The distribution of the phytoplankton size class (PSC) and the variations in the size classes are key to understanding ocean biogeochemical processes and ecosystem. Remote sensing of the PSC in the East China Sea (ECS) remains a challenge, although many PSC algorithms have been developed. Here based on a local dataset from the ECS, a regional model was tuned to infer the PSC from the spectral features of normalized phytoplankton absorption (a_{ph}) using a principal component analysis approach. Before applying the refined model to the real MODIS (Moderate Resolution Imaging Spectroradiometer) data, reconstructing satellite R_{rs} at 412 and 443 nm becomes critical through modeling them from R_{rs} between 469 and 555 nm using multiple regression analyses. Satellite-derived PSC values compare well with those derived from pigment composition, which demonstrates the potential of satellite ocean color data to estimate PSC distributions in the ECS from space. The refined model was applied to a_{ph} derived from R_{rs} observations collected by MODIS over the ECS from 2003 to 2016. Seasonal images show that the PSC distribution was heterogeneous in both temporal and spatial scales. Seasonal variations of the PSC in the ECS were probably affected by a combination of the water column stability, upwelling, sea surface temperature, and the Kuroshio Current. Additionally, human activity and riverine discharge may also influence the PSC distributions in the ECS, especially in coastal regions.

1. Introduction

Phytoplankton size class (PSC) is fundamentally important for ocean biogeochemical processes and ecosystems, especially for photosynthesis efficiency (Bouman et al., 2005; Uitz et al., 2008), primary production, and the transport of carbon (Kiørboe, 1993; Guidi et al., 2009; Hirawake et al., 2011). Thus, knowledge of the PSC dynamics can contribute to the improvement of our understanding of marine ecological and biogeochemical cycles. The classical size fractions of phytoplankton proposed by Sieburth et al. (Sieburth et al., 1978) include: micro- (>20 μm), nano- (2-20 μm), and pico-phytoplankton (<2 μm). Among the methods to measure PSC from water samples, which include microscopy (Montagnes et al., 1994), Coulter counter method (Sheldon and Parsons, 1967), and flow cytometry (Sun et al., 2000), pigment concentrations by high-performance liquid chromatography (HPLC) is the most systematic and qualify-controlled method (Van Heukelem and Hooker, 2011). However, these methods are time-consuming and methodologically complex. Furthermore, the large spatial and temporal variabilities make it difficult to continuously monitor the PSC using the field sampling methods.

Realistically, satellite ocean color data can provide synoptic observations, which are ideal for investigating the PSC at large spatial and temporal scales. In recent years, various algorithms have been designed to estimate the PSC using *in situ* data and ocean color data on both global and regional scales (IOCCG, 2014). Mostly algorithms can be partitioned into two categories, namely, "abundance-based" and "spectral-based" methods. The "abundance-based" methods are based on the statistical relationship between the PSC and phytoplankton abundance using measurements such as chlorophyll-a concentrations (Chla) (refer to Bracher et al. (2017) Table 2). These approaches rely on the assumption that high and low Chla waters are dominated by large and small phytoplankton, respectively. The "spectral-based" methods utilize the relationship between the variations in inherent optical properties with changes in the PSC using measurements such as phytoplankton absorption (a_{ph}), remote sensing reflectance (R_{rs}), and particulate backscattering (b_{bp}) (refer to Bracher et al. (2017) Table 2).

The East China Sea (ECS) is the base of the marine fishery resources in China and is one of the most productive ocean areas in the world (Furuya et al., 1996). Ascertaining the distribution of PSC can provide valuable information on the state of marine ecosystem and primary production in the area. Recent efforts have been focused on investigating the phytoplankton community

and size classes in the ECS and have suggested that the PSC displays obvious spatiotemporal heterogeneity in this region (Li et al., 2007; Luan et al., 2007; Jiang et al., 2014). For instance, Chen (2000) investigated the PSC and primary productivity in the marginal regions of the southern ECS using field data. The results showed that the PSC and their contributions to primary production displayed significant spatial differences in the shelf waters, upwelling waters, and Kuroshio water. Furuya et al.

5 (2003) presented the phytoplankton dynamics in the ECS in the spring of 1994 and the summer of 1996 using HPLC-derived pigment signatures. They found that a distinct horizontal heterogeneity in the phytoplankton composition was observed in the spring, and a "two-layer" distribution of phytoplankton appeared both off and on the shelf in the summer. Liu et al. (2016) used 7-years (2006-2012) field measurements to investigate the seasonal and spatial variations of major phytoplankton groups in the ECS, and found that monsoon forcing was a key factor to impact phytoplankton dynamics at seasonal scale.

10 Note that previous investigations on the PSC in the ECS have been conducted based on field observations, which may not reflect the real variation patterns of PSC. To our knowledge, no study has attempted to examine PSC distributions in the ECS at synoptic scales from satellite observations. Consequently, the dynamics of the PSC at different spatial and temporal scales and their mechanisms in the ECS are still poorly understood. In the ECS, Wang et al. (2014) found that the correlation between the variation patterns of the PSC and total Chla is not valid, and pointed out that the "abundance-based" methods for estimation

15 of PSC are probably not applicable in the ECS. Therefore, Wang et al. (2015) proposed an algorithm to model the PSC in the ECS using the spectral shape of normalized $a_{ph}(\lambda)$ through principal component analysis (PCA). This model shows good performance for estimating the PSC from both *in situ* a_{ph} and remote sensing reflectance (R_{rs}). However, note that this model was developed based on *in situ* data mainly from offshore waters of the ECS and off the coastal Japan; more importantly, the Wang et al. (2015) model has not been implemented in real satellite data in that study.

20 Therefore, the goals of this study were to: (1) refine the Wang et al. (2015) model for regional application in the ECS using an extensive dataset that includes samples covering highly varied waters and various seasons, (2) apply the refined PSC model to real Moderate Resolution Imaging Spectroradiometer (MODIS) satellite data, and (3) then preliminarily investigate previously unknown seasonal and spatial variation patterns of the PSC in the ECS.

2. Materials and methods

2.1 Study Area and Sampling stations

The East China Sea is one of the largest marginal seas in the western North Pacific and is bounded by China, Korea, and Japan (Fig. 1a). Nearly 70% of the ECS is occupied by a continental shelf shallower than 200 m. Numerous rivers flow into the ECS

5 from mainland China, including the Changjiang River which provides nearly 90% of the riverine discharge to the ECS (Zhang et al., 2007). In addition, the ECS experiences strong currents and multiple water masses, such as Changjiang diluted water (CDW), shelf mixed water, and the Kuroshio Current (Su and Yuan, 2005; Wang et al., 2014). Because of highly variable environmental conditions, the ECS exhibits complex marine biogeochemical processes and ecosystem.

The field measurements used in this study were collected from approximately 10 cruises over the last decade. These sampling

10 stations were distributed irregularly in the ECS and a few were in the Tsushima Strait (Fig. 1a). The field dataset encompassed various seasons and environmental conditions of the ocean, including turbid waters in the mouth area of Changjiang river, less turbid coastal water, and clear water away from the coast and in the open ocean. This field dataset consisted of *in situ* $a_{ph}(\lambda)$ and $R_{rs}(\lambda)$ data and phytoplankton pigment concentrations measured by HPLC. In total, 69 samples with synchronous measurements of pigments, a_{ph}, and R_{rs} data, 101 samples with coincident pigments and *in situ* a_{ph} data, and 27 samples with

15 only measured R_{rs} were available, and Fig. 1a shows the spatial distribution of samples.

The Kuroshio water in our study area suffers from a paucity of *in situ* R_{rs} data. Hence, in addition to the regional dataset, 201 *in situ* R_{rs} data points collected in the North Pacific and North Atlantic oceans (Fig. 1b) from the NASA SeaBASS archive were used as a supplementary dataset. The SeaBASS dataset was only used for algorithm development to reconstruct satellite R_{rs} data, along with our regional field dataset (see Section 2.4). The mean spectral shapes of the 14-years (2003-2016) MODIS

20 R_{rs} data in the North Pacific and North Atlantic oceans were similar to that in the Kuroshio water (Fig. 1c). Thus, *in situ* R_{rs} data collected in the North Pacific and North Atlantic oceans were used in the present study, although distribution regions of these data were beyond our study area.

2.2 *In situ* Measurements

Surface water samples (0-3 m) were collected with Niskin samplers mounted on a CTD Rosette or a clean bucket. These water samples were used for measurements of $a_{ph}(\lambda)$ and pigment concentrations.

2.2.1 Measurement and analysis of HPLC-derived PSC

5 For pigment analysis, seawater samples were filtered onto 47 mm Whatman GF/F glass fiber filters under gentle pressure (<0.01 MPa), and then stored initially on board in liquid nitrogen (-70 ℃) for later analysis in the laboratory. Briefly, the concentrations of 19 pigments were determined by reverse-phase HPLC following Van Heukelem and Thomas (2001). In our study, the diagnostic pigment analysis (DPA) approach was used to compute the PSC values from HPLC pigment data (hereafter referred to as the HPLC-derived PSC). The DPA was improved by Uitz et al. (2006) based on Vidussi et al. (2001) method. In brief, the DPA uses seven diagnostic pigment concentrations to obtain the HPLC-derived PSC, including

10 fucoxanthin (C_f), peridinin (C_p), 19′-hexanoyloxyfucoxanthin (C_h), 19′-butanoyloxyfucoxanthin (C_b), alloxanthin (C_a), Chlorophyll-b (C_{Cb}), and zeaxanthin (C_z). The HPLC-derived PSC values were then given by:

$$f_{micro} = \left(1.41C_f + 1.41C_p\right) / \sum W_i P_i \tag{1}$$

$$f_{nano} = \left(0.60C_a + 0.35C_b + 1.01C_{Cb} + x \times 1.27C_h\right) / \sum W_i P_i \tag{2}$$

$$f_{pico} = \left(0.86C_z + y \times 1.27C_h\right) / \sum W_i P_i \tag{3}$$

15 where f_{micro}, f_{nano}, and f_{pico} denote the size fractions of micro-, nano-, and pico-phytoplankton, respectively. x and y are the proportions of nano- and pico-phytoplankton in Hex, respectively (see Brewin et al. (2010) for more details). $\sum W_i P_i$ is the weighted sum of the seven diagnostic pigments, which is defined as (Uitz et al., 2006):

$$\sum W_i P_i = 1.41C_f + 1.41C_p + 0.60C_a + 0.35C_b + 1.27C_h + 0.86C_z + 1.01C_{Cb} \tag{4}$$

2.2.2 Measurement of a_{ph}

20 To obtain a_{ph} data, we used the quantitative filter technique (QFT) via a series of processes (Mitchell, 1990). Water samples were filtered through 25 mm Whatman GF/F glass fiber filters under gentle pressure, and immediately frozen on board in

liquid nitrogen. In our study, the "transmittance" approach was used for the samples that were collected from southern Jeju

Island and the Tsushima Strait (hereafter referred to as dataset-1). The optical density (OD) values of total particles were

measured using a dual-beam multi-purpose spectrophotometer between 350 nm and 750 nm at 1 nm resolution. Similarly, we

measured the OD value of the detritus after extracting the phytoplankton pigments in methanol at least 24 h. Meanwhile, a

5   blank filter saturated with pure seawater was used as the reference filter. Then, the absorption coefficients of total particles

$a_p(\lambda)$ and detritus $a_d(\lambda)$ were calculated from the corresponding OD values based on a correction of Cleveland and Weidemann

(1993). In contrast, the "transmittance-reflectance" approach was performed on the samples that were collected from the

coastal and offshore regions of Zhejiang and Jiangsu (hereafter referred to as dataset-2). The optical densities of total particles,

detritus, and the reference filter were obtained in both "transmission mode" and "reflection mode" between 250-850 nm at 1

10   nm resolution using a PerkinElmer lamda650s. Then, we converted these OD values into $a_p(\lambda)$ and $a_d(\lambda)$ values using the

method of Tassan and Ferrari (1995; 2002). Finally, the a_{ph} data were obtained as the difference of $a_p(\lambda)$ - $a_d(\lambda)$ at all sampling

stations.

2.2.3 Measurement of R_{rs}

To obtain R_{rs} data in dataset-1, the PRR-800/811 was used to measure the vertical profiles of the downwelling irradiance $E_d(\lambda,z)$

15   and upwelling radiance $L_u(\lambda,z)$ at 13 spectral channels (380, 412, 443, 465, 490, 510, 532, 555, 565, 589, 625, 665, and 683

nm). The water-leaving radiance $L_w(\lambda)$ was then determined from the profile of $L_u(\lambda,z)$ (Hirawake et al., 2011). The above-

water surface downwelling irradiance $E_d(\lambda,0^+)$ was simultaneously measured by a cosine collector. Then, $R_{rs}(\lambda)$ data were

calculated as the ratio of $L_w(\lambda)$ to $E_d(\lambda,0^+)$. For the purpose of consistency with satellite observations that characterize the

oceanic surface layer, our analysis exclusively considered the near-surface R_{rs} data.

20   For dataset-2, R_{rs} data was collected under suitable solar illumination (generally between 9am and 3pm local time) using an

ASD FieldSpec spectroradiometer in the spectral range of 350-1050 nm with 1.5 nm increments. The radiance spectra of water,

sky, and a gray reference panel were measured following the above-water measurement approach (Mueller et al., 2003). For

each of the three targets, ten spectra were collected and then averaged after removing abnormal spectra. According to the

Ocean Optics Protocol (Mueller et al., 2003), the $R_{rs}(\lambda)$ data were obtained as

$$R_{rs}(\lambda) = \left(L_t - r * L_{sky}\right) / \left(L_p * \pi / \rho_p\right) \qquad (5)$$

where L_t, L_{sky}, and L_p correspond to the radiance values measured from the water, sky, and the reference panel, respectively. ρ_p is the diffuse reflectance of the reference panel provided by the manufacturer. γ is the surface Fresnel reflectance related to the

5 wind speed (2.6% - 2.8% for 10 m s^{-1} wind, 2.5% for <5 m s^{-1} wind, 2.2% for calm weather) (Tang et al., 2004).

All R_{rs} and a_{ph} data were resampled at the centers of MODIS wavebands (412, 443, 469, 488, 531, 547, 555, 645, 667, and 678 nm) using the spectral response function of the MODIS sensor.

2.3 Satellite data

The global MODIS Level 3 standard monthly R_{rs}, Chla, and sea surface temperature (SST) products (about 4 km resolution)

10 from 2003 to 2016 were provided by the NASA Ocean Color website (http://oceancolor.gsfc.nasa.gov/). The dataset corresponding to our study area (25-35° N and 118-132° E) was extracted from these global coverage datasets. These regional R_{rs} products were processed using the MathWorks MATLAB software to obtain the satellite-derived PSC. Additionally, daily Level 2 R_{rs} data (1 km resolution) from MODIS sensor were downloaded from the NASA Ocean Color website.

Samples were matched to daily R_{rs} data to assess the accuracy of satellite-derived a_{ph} and PSC results. To ensure the validity

15 of satellite data before the matchup analysis, the following constraints were applied to the matchup dataset: (1) the matchup dataset only included satellite data with an overpass time window within 5 h before and after the field measurements; (2) to reduce the effect of outliers, median R_{rs} value for a 3×3 window centered on the sampling station coordinates was defined as satellite R_{rs} data; (3) negative MODIS R_{rs} data were eliminated from the matchup analysis. Based on these criterion, 21 satellite matchups with coincident measured R_{rs}, and 23 satellite matchups with coincident measured PSC and a_{ph} were available, as

20 shown in Fig. 1a.

2.4 Model accuracy assessment

To evaluate the consistency between the derived and measured values, the Pearson correlation coefficient (R), root mean square

error (RMSE), and mean absolute percentage error (MAPE) were used. Statistical assessments were performed in $\log_{10}$ space

for the phytoplankton absorption coefficient and in linear space for the phytoplankton size class. These statistical indicators

can be written as:

$$RMSE = \frac{1}{n}\sqrt{\sum_{i=1}^{n}\left[\left(x_{i,\text{derived}} - x_{i,\text{field}}\right)/x_{i,\text{field}}\right]^2} \qquad (6)$$

$$MAPE\left(\%\right) = \frac{1}{n}\sum_{i=1}^{n}\left|\left(x_{i,\text{derived}} - x_{i,\text{field}}\right)/x_{i,\text{field}}\right| \times 100\% \qquad (7)$$

where n is the number of samples. $x_{i,\text{derived}}$ and $x_{i,\text{field}}$ are the derived and measured data for the i-th sampling station, respectively.

2.5 Modified Wang et al. (2015) model for retrieving PSC

Wang et al. (2015) developed an spectral-based PSC model to quantify the size fraction of three phytoplankton classes using

the spectral shape of $a_{\text{ph}}(\lambda)$ through PCA approach. Details of the development and parameterization of the model are described

in Wang et al. (2015). In brief, to reduce the biomass effects, the normalized $a_{\text{ph}}(\lambda)$ (hereafter referred to as $a_{\text{ph}}^{\text{std}}(\lambda)$) was

computed by the ratio of $a_{\text{ph}}(\lambda)$ to their wavelength mean values. Then, PCA was applied to the $a_{\text{ph}}^{\text{std}}(\lambda)$ data to capture the

spectral variation in the phytoplankton absorption related to cell size. The input for PCA is a $m \times N$ matrix constituted of $a_{\text{ph}}^{\text{std}}$

(λ), where m and N are the number of input wavelengths and samples, respectively. The output of PCA comprised two terms,

principal component (PC) scores and PC weights (also called loading factors). The PC scores were assumed to correlate with

the size classes. Therefore, the relationship between the size fraction of micro- or pico-phytoplankton and PC scores was

established using a logistic-type regression model (Hosmer Jr et al., 2013), as follows:

$$f_t = 1/\left[1 + \exp\left(-\beta_0 - \sum_{i=1}^{k}\beta_i S_i\right)\right], \quad S_i = \sum_{j=1}^{m} w_{ij}a_{\text{ph}}^{\text{std}}(\lambda_j) \qquad (8)$$

where f_t denotes the size fraction of phytoplankton class (t = micro or pico). β_0 and β_i are the regression coefficients between

f_t and PC scores. k is the number of PC scores (k = 4 in this study). w_{ij} refers to the loading factor for the i-th PC. m is the

number of wavelengths. Similar to previous studies (Brewin et al., 2010; Hirata et al., 2011), f_{nano} was calculated as $1 - f_{\text{micro}} - f_{\text{pico}}$, by considering that the sum of the three phytoplankton size fractions was 1.

The $a_{ph}(\lambda)$ at the MODIS bands were derived from $R_{rs}(\lambda)$ data using the quasi-analytical algorithm (QAA) proposed by Lee et al. (2002). QAA was used in this study because it does not suppose a fixed shape for $a_{ph}(\lambda)$ (Lee et al., 2002, 2009). Because QAA could give satisfactory retrievals of $a_{ph}(\lambda)$ at the first 6 MODIS bands (i.e., 412, 443, 469, 488, 531, and 547 nm), as shown later for details, only $a_{ph}(\lambda)$ data at these bands were used for the PSC model development in this study (i.e., $m=6$ in

5   Eq.(8)).

Before QAA approach and the refined PSC model were applied to real MODIS data, the reconstruction approach was used to reconstruct satellite $R_{rs}(\lambda)$ at 412 and 443 nm. Satellite $R_{rs}(412)$ and $R_{rs}(443)$ were quantified as multivariable linear regression model using R_{rs} data from 469 to 555 nm and regression coefficients (K_0 and K_i) as follows:

$$R_{rs}^{rc}(\lambda) = \sum_{i=1}^{n} K_i R_{rs}(\lambda_i) + K_0 \tag{9}$$

10   where $R_{rs}^{rc}(\lambda)$ is the reconstructed R_{rs} data at wavelength λ (λ = 412 or 443 nm); $R_{rs}(\lambda_i)$ is the input R_{rs} value at five MODIS wavelengths (λ_i = 469, 488, 531, 547, and 555 nm); K_0 and K_i are the coefficients determined from multivariant regression.

3 Result

3.1 Regional tuning of the PSC model for the ECS

Following Wang et al. (2015), Eq. (8) was fitted to the 170 *in situ* a_{ph} and HPLC-derived PSC data using a non-linear least

15   square fitting procedure for developing the PSC model. The established parameters and associated R and RMSE values for each of the fits are shown in Table 1. Fig. 2 shows the strong linear relationships between the *in situ* a_{ph}^{std}-derived PSC and HPLC-derived results, with R values of 0.89, 0.70, and 0.84 and RMSE values of 0.11, 0.11, and 0.11 for micro-, nano-, and pico-phytoplankton, respectively. The samples were close to the 1:1 line, with most of the samples within the ± 20% fraction range.

20   We also examined the feasibility of the PSC model for satellite observations by coupling QAA, using R_{rs} data coincident with HPLC-derived PSC and *in situ* a_{ph}. First, we used QAA version 5 (QAA_v5) to yield a_{ph} from measured R_{rs}. To assess the performance of QAA_v5, negative retrieved a_{ph} values were eliminated, and the remainder were compared with the measured

values at all MODIS bands (Fig. 3). The derived a_{ph} values by QAA_v5 show reasonably good agreement with the *in situ* a_{ph} values at short wavelengths from 412 to 547 nm, with high R values and low RMSE and MAPE values. In contrast, the performance of QAA_v5 was poor and produced large overestimation of a_{ph} at long wavelengths, especially at 645, 667, and 678 nm. These results were similar to previous findings by Lee et al. (2014) and Tiwari et al. (2014). These results clearly

5   demonstrate that QAA_v5 can produce accurate a_{ph} values at 412, 443, 469, 488, 531, and 547 nm. Therefore, as previously stated, only a_{ph} values at these bands were used to calibrate the PSC model. Then, the PSC values were inferred from the retrieved a_{ph} using Eq. (8) with the established parameterizations. As shown in Fig. 4, the QAA $a_{ph}^{std}(\lambda)$-derived PSC values were consistent with HPLC-derived results, and almost all of the points fell within the ± 20% fraction range. The R and RMSE values were 0.79 and 0.13 for micro-phytoplankton, 0.43 and 0.12 for nano-phytoplankton, and 0.80 and 0.13 for pico-

10   phytoplankton. These results suggest that the refined PSC model for the ECS coupling QAA_v5 is able to accurately estimate the PSC from remote sensing reflectance R_{rs}.

3.2 Comparison of satellite R_{rs} with *in situ* measurements

Before applying the PSC model to real MODIS data, we assess the accuracy of MODIS R_{rs} based on the synchronous *in situ* measurements. Table 2 shows the statistical results of the comparison between satellite R_{rs} and *in situ* values for MODIS bands.

15   For MODIS R_{rs} data, we found a reasonably good consistency at green and red bands (from 469 to 555 nm), with R values within 0.85-0.97 and MAPE values within 14.90% - 27.25%. Although the R values were > 0.85, the MAPE values were > 54% at 645, 667, and 678 nm. This is probably caused by the lower R_{rs} values at these bands due to strong absorption of water itself. In addition, a low accuracy was observed at 412 and 443 nm. The R values were 0.46 and 0.73, and the MAPE values were 47.33% and 36.90% at 412 and 443 nm, respectively. The high noise and low accuracy at these two bands are suggested

20   to be caused by the uncertainty of the atmospheric correction procedures and significant band degradation (Meister, 2011; Hu et al., 2013). Considering the importance of $R_{rs}(412)$ and $R_{rs}(443)$ to QAA algorithm, the poor accuracy of satellite R_{rs} at these bands may introduce uncertainty in the retrieved a_{ph} data, which further increases the uncertainty of satellite-derived PSC. Thus, an accurate assessment of satellite-derived PSC requires the improved quality of satellite $R_{rs}(412)$ and $R_{rs}(443)$. In this

study, the reconstruction approach (Eq. (9)) was used to fulfill this objective.

3.3 Reconstruction of MODIS R_{rs} data

The reconstruction function (Eq. (9)) was applied to the regional field dataset and SeaBass dataset to obtain the regression

coefficients. The resulting relationships between the *in situ* and modeled R_{rs} show strong agreement, with high R^2 and low

5   MAPE values (Fig. 5). For 412 and 443 nm, the R^2 values were close to 1.0 with a significance level of *p*-value <0.01. Their

MAPE values were both lower than 9.0%. The multivariant functions with established coefficients were applied to the original

MODIS R_{rs} data to obtain the reconstructed satellite $R_{rs}^{rc}(412)$ and $R_{rs}^{rc}(443)$ data. Table 3 shows the comparison of the original

satellite R_{rs} and the satellite R_{rs}^{rc} data with measured R_{rs} at 412 and 443 nm. From Table 3, it is observed that the satellite R_{rs}^{rc}

data were in better agreement with the *in situ* R_{rs} data than the original satellite R_{rs}, especially at 412 nm. At 412 nm, the values

10   of R, RMSE, and MAPE reached 0.70, 0.0019, and 35.15% after reconstruction, respectively, while these values were 0.46,

0.0026, and 47.33% for the original satellite data, respectively. These results indicated that the accuracy of the satellite R_{rs} data

at 412 and 443 nm could be improved through reconstruction using the selected MODIS bands.

3.4 Validation of satellite-derived absorption coefficient and PSC with *in situ* data

Based on the above analysis, we used the satellite $R_{rs}^{rc}(412)$ and $R_{rs}^{rc}(443)$ data rather than the original satellite R_{rs} data to compute

15   a_{ph} using QAA_v5. Fig. 6 shows the comparison of the a_{ph} data derived from satellite R_{rs}^{rc} (denoted by a_{ph}^{rc}) and a_{ph} derived from

original satellite R_{rs} with *in situ* measurements at the six MODIS bands. Table 4 summarized their corresponding statistical

comparisons, i.e., R, RMSE, MAPE and percentage of valid points (PVP). Here PVP is defined as the ratio of the number of

positive satellite-derived values (n) to the total number of matchups (N) (as PVP = $n/N \times 100\%$). For the satellite-derived a_{ph}^{rc},

the R values were above 0.80 ($R = 0.69$ at 547 nm), and were significantly higher than those for the satellite-derived a_{ph} (with

20   most of the values below 0.6). The statistics (RMSE and MAPE) the satellite-derived a_{ph}^{rc} are also generally better than those

for the satellite-derived a_{ph}. In addition, compared with satellite-derived a_{ph}, the PVP for the satellite-derived a_{ph}^{rc} significantly

increased with an average of 23.48%. Meanwhile, Fig. 6 also shows that the satellite-derived a_{ph}^{rc} had more valid samples and

were more clustered around the 1:1 line than the satellite-derived a_{ph}. Overall, both Table 4 and Fig. 6 indicated that the satellite-derived a_{ph} had poor accuracy and low PVP values, whereas the accuracy of satellite-derived a_{ph}^{rc} can be significantly improved with more valid samples through the reconstruction of satellite R_{rs} data.

We applied the refined PSC model to the a_{ph} data derived from the original satellite R_{rs} to estimate the satellite R_{rs}-derived PSC

5   (Fig. 7a-c). It can be seen that the satellite R_{rs}-derived PSC were inconsistent with the HPLC-derived results, showing obvious under and overestimations of the retrieved PSC for most of the samples. Their R values were all below 0.27. For comparison, we also estimated the PSC from the satellite a_{ph}^{rc} derived from the reconstructed R_{rs}^{rc} data, denoted as the satellite R_{rs}^{rc}-derived PSC. Fig. 7d-f show the comparison of the satellite R_{rs}^{rc}-derived PSC with the HPLC-derived values. The satellite R_{rs}^{rc}-derived PSC values agreed well with the HPLC-derived results. The R values of 0.68, 0.46, and 0.64 and RMSE values of 0.13, 0.13,

10   and 0.19 were observed for micro-, nano-, and pico- phytoplankton, respectively. For micro-, nano-, and pico-phytoplankton, almost all of the samples fell within the ±20% fraction ranges, although a slight underestimation of f_{pico} occurred in a few samples. These results indicated that the use of satellite R_{rs}^{rc} could significantly improve the performance of the refined PSC model on satellite observations, which result in satisfactory satellite-derived PSC results with acceptable errors. Therefore, we further investigated the spatiotemporal variability of the PSC in the ECS based on satellite-derived products.

15   **3.5 Seasonal patterns of the PSC distribution in the East China Sea**

To describe the large-scale seasonal variability of the PSC in the ECS, the PSC method was applied to 14 years (2003-2016) of MODIS monthly R_{rs} data to obtain monthly PSC products. Then, seasonal composite PSC images were generated by averaging the monthly PSC products over a three month period for each season. In this study, spring, summer, autumn, and winter were defined as March to May, June to August, September to November, and December to February of the next year,

20   respectively.

Fig. 8 depicted the seasonal variation patterns of the PSC in the ECS in the four seasons, which indicated that the size fractions of three phytoplankton classes were heterogeneous in both temporal and spatial scales. In the spring (Fig. 8a-c), the f_{micro} values were higher (0.45-0.90) on the ECS shelf sea than those in the Kuroshio water (<0.3). The highest f_{micro} were found in coastal

waters. Relatively high f_{nano} (0.4-0.6) were clearly observed on offshore shelf and in southern Japan. However, pico-phytoplankton were the dominant size class (f_{pico} = 0.5-0.75) over the southeastern ECS. During the summer (Fig. 8d-f), the size fractions of micro-phytoplankton were high (0.45-0.85) in coastal waters. The high f_{micro} tongue-shape structure near the Changjiang Bank extended toward southeast along the 30 m isobaths. High nano-phytoplankton proportions occurred in the

5 ECS shelf sea with water depths of 30–200 m. The pico-phytoplankton contributions to Chla were relatively high around the ECS shelf break. Pico-phytoplankton (f_{pico} >0.6) represented the most abundant size class in the areas deeper than 200 m, which was similar to the results from the field measurements by Chen (2000). In the autumn (Fig. 8g-i), high f_{micro} were found in coastal waters and extended over the area shallower than 50 m isobaths. The proportion patterns of nano- and pico-phytoplankton in the autumn were broadly similar to those in the summer. However, high nano- phytoplankton proportions

10 were also in the northern Japan. In the winter (Fig. 8j-l), high f_{micro} were mainly distributed on the ECS shelf. The regions with higher f_{micro} (>0.7) extended outward, and connected to the area around the Korean coast. The distributions of the size fractions of nano- and pico-phytoplankton were broadly similar to those in the spring. For nano-phytoplankton, relatively high f_{nano} were also observed in some areas of the Kuroshio water in the winter.

3.6 Area difference in the monthly climatological PSC in the ECS

15 Since the East China Sea is so extensive, with a number of different environmental conditions and ecosystems, specific regions were selected for further investigation. Fig. 9a shows the locations of the selected subareas, which were the mouth area of Changjiang river (MCJR), middle shelf region (MSR), and Kuroshio region (KR). These subareas were selected based on geographical region and physical driving force. With each of the subareas, this study investigated averages of the monthly climatological PSC, as well as Chla concentration (Fig. 9 b-d).

20 In the MCJR (Fig. 9b), micro-phytoplankton comprised 60-80% of the Chla throughout the year, with the maximum value found in April and relatively low fraction in the summer-early autumn (June-September). Nano-phytoplankton comprised 18-30% of the Chla, while the contribution of pico-phytoplankton to Chla was below 10% throughout the year. In the MSR (Fig. 9c), the micro-phytoplankton proportions were greater than that of pico-phytoplankton during the winter and spring, while the

opposite was found in the summer and autumn. The pico-phytoplankton in the MSR was highest in August and September, with a peak in the summer and early autumn (June-September). Nano-phytoplankton were dominant (40-50%) for most of the year in this region. In contrast, the Kuroshio region domain showed a predominance of pico-phytoplankton (50-90%) throughout the year, except in January and February, with higher proportions observed in the summer (Fig. 9d). The nano-

5 phytoplankton proportions were slightly lower than pico-phytoplankton in the winter, however, their proportions were low (<30%) in the rest of the year. The size fractions of micro-phytoplankton in the KR remained below 18% throughout the year, except in April (23%) (Fig. 9d).

4 Discussion

4.1 Satellite application of the refined PSC model

10 The most important advantage of satellite ocean color data is the ability to provide information on the spatiotemporal variability of the PSC. However, remote sensing of the PSC in the ECS is still a challenge task, although many "abundance-based" and "spectral-based" algorithms have been designed using field measurements and satellite data in the world's oceans. Taking the optical property in the ECS, Wang et al. (2014) reported that the "abundance-based" approaches are not necessarily applicable in the ECS, and the absorption spectra of phytoplankton could instead be used to obtain the PSC in the ECS. More than 80%

15 of the variability in the spectral shape of phytoplankton absorption was highly related to the changes in the size classes (Ciotti et al., 2002; Bricaud et al., 2004). Therefore, in this study we regionally tuned the Wang et al.(2015) model for deriving PSC in the ECS from the spectral variation of a_{ph}. However, the application of this PSC model to original MODIS data has hampered, as showed in this study (Fig. 7a-c). This may be related to the low accuracy of the MODIS R_{rs} at 412 and 443 nm (Table 2), which introduced additional uncertainties into the satellite-derived a_{ph} from R_{rs} (Fig. 6; Table 4). Because the retrieved accuracy

20 of the PSC depends heavily on the retrieval accuracy of $a_{\mathrm{ph}}(\lambda)$, the uncertainties of satellite-derived a_{ph} may lead to a very large uncertainty in the satellite-derived PSC (Fig. 7a-c). To solve this problem, the multivariant function was employed to reconstructed MODIS $R_{\mathrm{rs}}(412)$ and $R_{\mathrm{rs}}(443)$ values using satellite R_{rs} from 469 to 555 nm. Previous studies have reported that

the use of multiple band spectral bands could successfully reconstruct hyperspectral R_{rs} data (Lee et al., 2014; Sun et al., 2015).

In our study, the use of satellite R_{rs}^{rc} improved the accuracy and PVP of the satellite-derived a_{ph}^{rc} data using QAA_v5 (Fig. 6; Table 4), and dramatically improved the accuracy of the satellite-derived PSC (Fig. 7d-f). The R and RMSE for all size fractions derived from the reconstructed satellite data were 0.7 and 0.15 respectively, as compared to the values of 0.064 and 0.38

5   respectively obtained from original satellite data. Overall, this study successfully estimated the PSC in the ECS from MODIS remote sensing reflectance. The findings presented here complement and enhance recent studies that have demonstrated that satellite ocean color data can be used to retrieve the PSC in the ECS. Although reconstruction of satellite R_{rs} data and the refined PSC model appear promising, the need for improved availability of *in situ* R_{rs} data covering various hydrological and trophic conditions in the ECS (e.g., the Kuroshio water) to ultimately improve the PSC estimations from satellite data. It is

10   also clear that because the refined PSC model are based on data collected from the ECS, its application to other areas may be limited and require further investigations based on *in situ* and satellite data beyond the ECS.

4.2 Seasonal variations of the PSC in the ECS

The seasonal distributions of the PSC in the ECS and the monthly climatological PSC in the subareas (Fig. 8 and Fig. 9) clearly showed that the PSC varied across both spatial and temporal scales in the ECS. In general, micro-phytoplankton are favoured

15   under environmental condition with stronger mixing and high-nutrient, while pico-phytoplankton tend to dominate in the low-nutrient waters (Stæhr et al., 2004; IOCCG, 2014; Lamont et al., 2018). Combined with the monthly climatological satellite-derived Chla data (Fig. 9), we discussed the related physical and biochemical effects to provide a better understanding of the temporal-spatial variations of the PSC in the East China Sea.

In the spring, the spring bloom occurred in the mouth area of Changjiang river and middle shelf region (Fig. 9 b and c). Weak

20   wind stress can retain the vertical mixing of the water column. Nutrients are transported to the upper layer from the nutrient-rich deep layers (Behrenfeld et al., 2006; Boyce et al., 2010). This enhanced nutrient condition results in higher supports phytoplankton biomass (Fig. 9 b and c) and favours the presence of micro-phytoplankton on the ECS shelf sea (Fig. 8a). This was consistent with previous studies showing that the large cell sizes, such as diatoms and *Prorocentrum donghaiense*, were

dominant on the ECS shelf sea in the spring (Furuya et al., 2003; Lou and Hu, 2014). Around the shelf break, upwelling

frequently occurs at this region on the ECS to the Kuroshio water (Chen et al., 2009), transposing nutrients-rich waters from

the subsurface layers to the euphotic zone. Thus, upwelling may be one of the factors for explaining the nano-phytoplankton

dominated communities around the shelf break in the spring (Fig. 8b). In contrast to the pattern on the ECS shelf sea, the

5   Kuroshio water exhibits relatively low Chla (<0.5 mg m^{-3}) (Fig. 9d), and the pico-phytoplankton were the dominant size class

in the southeast ECS shelf with water depths > 200 m in the spring (Fig. 8c). This may be related to the Kuroshio Current. The

mainstream of the Kuroshio Current strongly flows northeastward along around the 200 m isobath (Katoh et al., 2000), carrying

warm and low nutrients water (Jiao et al., 2005). The high surface temperature could increase the water column stability,

thereby diminishing the nutrient supply to the surface layer (Lovelock, 2007). Thus, this nutrient-depleted condition leads to

10   the low phytoplankton biomass and promotes the growth of pico-phytoplankton in this region. This has also been confirmed

by the positive correlation between SST and the pico-phytoplankton proportion in the KR (Table 5).

In the summer, due to surface warming and low wind stress, the reduced wind mixing and stronger thermal stratification results

in the less nutrient supply to the surface layers. Meanwhile, all nutrients are gradually consumed. In addition to surface heating,

the Kuroshio Current influences the stratification of the water column and further leads to oligotrophic conditions in the

15   Kuroshio water. Thus, the low Chla concentrations (<0.5 mg m^{-3}) were observed in the KR (Fig. 9d). The oligotrophic

conditions favour the presence of pico-phytoplankton as they have advantage in utilising the regenerated nutrients available in

the nutrient-depleted regions (Brewin et al., 2010; Barlow et al., 2016). Thus, the pico-phytoplankton were found to dominate

the low biomass areas deeper than 100 m (Fig. 8f). In striking contrast, micro-phytoplankton dominance was limited to the

inner part of the ECS shelf with shallower than 50 m isobath (Fig. 8d). This may be related to the enhanced nutrient

20   concentrations by river and estuarine discharge. Due to anthropogenic activities such as various agricultural and industry

activities, nutrient-rich water discharges into the East China Sea, especially in the summer (Siswanto et al., 2008; Zhang et al.,

2017). This is especially evident in the higher Chla concertation (> 6 mg m^{-3}) observed in the MCJR (Fig. 9b) in the summer.

The outer part of the ECS shelf was dominated by nano-phytoplankton (Fig. 8e and Fig. 9c), likely due to the enhanced nutrient

conditions in offshore waters resulting from the coastal region by the strong convection currents. The study of Yamaguchi et

al. (2012) revealed that the CDW takes approximately 2 months to reach the Tsushima Strait. Therefore, the nutrient supply

from riverine discharge may be a major controlling factor in the large cell sizes (micro- and nano- phytoplankton) in coastal

and offshore waters in the summer. Similarly, Jiang et al. (2015) reported that micro-sized diatoms and dinoflagellates

dominated the CJE and adjacent areas in the summer in response to the available nutrients.

5   When autumn arrives, as wind stress strengthens and temperature decreases, convectional mixing of the water column increases

and the stratification weakens, which bring nutrients upward from the lower layers. Thus, micro- and nano-phytoplankton

dominance was clearly observed in coastal waters and the ECS shelf sea (Fig. 8g-h and Fig. 9c). Conversely in the oligotrophic

Kuroshio water likely caused by the Kuroshio Current, pico-phytoplankton were still the dominant size class (Fig. 8i and Fig.

9d), and the Chla was still below 0.5 mg m^{-3} during this period (Fig. 9d).

10   In the winter, the stratification of the water column is broken down and vertical mixing develops, due to strong winds and sea

surface cooling. Nutrients are replenished from the underlying nutrient-rich waters, which favour the micro- and nano-

phytoplankton on the ECS shelf and some regions of the Kuroshio water (Fig. 8j-k). Intense mixing has also been reported to

promote larger phytoplankton cells (Taylor and Joint, 1990). For the oligotrophic Kuroshio water, the Chla concentration was

still low (Fig. 9d), and the pico-phytoplankton were relatively abundant (Fig. 8l and Fig. 9d) in the winter, which was similar

15   to that in other seasons.

4.3 Response of phytoplankton size class to sea surface temperature

It has previously been suggested that sea surface temperature is one of the important factors that influence the PSC dynamic

(Chen, 2000; Barnes et al., 2010; IOCCG, 2014). Fortunately, the SST parameter can be detected from remotely-sensed data.

It enables the study of the PSC response to the SST change under differing environmental conditions. Therefore, based on the

20   14-years (2003-2016) time series of the monthly SST and satellite-derived PSC data, we discussed the relationship between

SST and PSC in the three subareas of the ECS, which were proved with correlation analysis (Table 5). SST can influence the

vertical structure of the water column and would further change the nutrient supply and light conditions, which influence

phytoplankton size structure and community.

For the three subareas, all of the correlations between micro-phytoplankton proportion and SST were negative, especially for the middle shelf region (R= -0.76), suggesting that SST will restrict the growth of micro- phytoplankton. For nano-phytoplankton, the correlations between its size fractions and temperature were positive in the MCJR and MSR, but no significant correlation in the MSR, while the correlation was negative in the KR. It was reported that nano-phytoplankton (e.g, flagellates) were dominant in offshore in warmer shelf regions, because they are better utilising the increase in nutrient concentration after upwelled water has warmed (Barlow et al., 2016). In contract, significant positive correlations between pico-phytoplankton proportion and SST were identified in the three subareas. Similarly, Chen (2000) reported that there was a significant positive correlation between *in situ* pico-phytoplankton proportions and water temperature in the Kuroshio water. This may be because the increase (decrease) in density stratification due to higher (lower) SST may weaken (increase) strong mixing vertical, which would provide a favorable low-nutrient (high-nutrient) condition for pico- (micro-) phytoplankton. The result of Table 5 indicated the SST is a key factor that influence the PSC dynamic in the ECS. Additionally, the inter-annual variation of phytoplankton size classes in the ECS were complicated and could not be fully explained by the individual factor. Further research is still needed to understand the seasonal and inter-annual variabilities of the PSC and their responses to environmental factors (e.g., riverine discharge, monsoon forcing, and climate change) when more data become available.

5 Conclusions

In this study, the PSC model was regionally tuned for application to the ECS using extensive *in situ* data covering various seasons and environmental conditions in the ECS. When the refined model was applied to MODIS observations, there is a critical step to reconstruct satellite remote sensing reflectance at blue bands. It leads to reliable performance of the PSC model on real MODIS R_{rs} data, which showed a good agreement with the HPLC-derived PSC values, with almost all of the samples falling within ±20% - fraction range. Along the way, our present study preliminarily estimates spatial distributions of PSC in the ECS from space. The regional PSC model was applied to satellite data from MODIS during 2003 and 2016 to infer the PSC distribution at seasonal scales. The obtained results show that the PSC varied across both spatial and temporal scales in the ECS, which are consistent with prior knowledge of phytoplankton seasonal dynamics in the area. The seasonality of the

PSC in the ECS may be mainly related to the vertical structure of the water column, upwelling, sea surface temperature, and the Kuroshio Current. It is affected by riverine discharge and human activity, especially in coastal waters. However, the spatiotemporal distribution at different scales (monthly, seasonal and annual scales) and its mechanisms are needed to be investigate in details in the future.

Author Contribution

H. L. Zhang analysed the data and wrote the manuscript; S. Q. Wang and Z. F. Qiu contributed to design this study and interpretation of the results; D. Y. Sun and J. Ishizaka revised the draft; S. J. Sun and Y. J. He provided comments and suggestions to improve the manuscript.

Conflict of Interest

The authors declare that there has no conflict of interest.

Acknowledgements

This research was jointly supported by the National Key Research and Development Program of China (2016YFC1400904), the National Natural Science Foundation of China (41506200, 41576172, and 41276186), the Natural Science Foundation of the Jiangsu Higher Education Institutions of China (15KJB170015), the Provincial Natural Science Foundation of Jiangsu in China (BK20150914, BK20151526, and BK20161532), the National Program on Global Change and Air-sea Interaction (GASI-03-03-01-01), the Public Science and Technology Research Funds Projects of Ocean (201005030), a project funded by "the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD)", the Research and Innovation Project for College Graduates of Jiangsu Province (KYLX16_0952), and the China Scholarship Council.

References

Barlow, R., Gibberd, M., Lamont, T., Aiken, J., and Holligan, P.: Chemotaxonomic phytoplankton patterns on the eastern

boundary of the Atlantic Ocean, Deep Sea Research Part I: Oceanographic Research Papers, 111, 73-78, 2016.

Barnes, C., Irigoien, X., De Oliveira, J. A., Maxwell, D., and Jennings, S.: Predicting marine phytoplankton community size structure from empirical relationships with remotely sensed variables, Journal of Plankton Research, 33, 13-24, 2010.

Behrenfeld, M. J., O'Malley, R. T., Siegel, D. A., McClain, C. R., Sarmiento, J. L., Feldman, G. C., Milligan, A. J., Falkowski, P. G., Letelier, R. M., and Boss, E. S.: Climate-driven trends in contemporary ocean productivity, Nature, 444, 752-755, 2006.

Bouman, H., Platt, T., Sathyendranath, S., and Stuart, V.: Dependence of light-saturated photosynthesis on temperature and community structure, Deep Sea Research Part I: Oceanographic Research Papers, 52, 1284-1299, 2005.

Boyce, D. G., Lewis, M. R., and Worm, B.: Global phytoplankton decline over the past century, Nature, 466, 591-596, 2010.

Bracher, A., Bouman, H. A., Brewin, R. J., Bricaud, A., Brotas, V., Ciotti, A. M., Clementson, L., Devred, E., Di Cicco, A., and Dutkiewicz, S.: Obtaining phytoplankton diversity from ocean color: a scientific roadmap for future development, Frontiers in Marine Science, 4, 55, 2017.

Brewin, R. J., Sathyendranath, S., Hirata, T., Lavender, S. J., Barciela, R. M., and Hardman-Mountford, N. J.: A three-component model of phytoplankton size class for the Atlantic Ocean, Ecological Modelling, 221, 1472-1483, 2010.

Bricaud, A., Claustre, H., Eacute, Ras, J., and Oubelkheir, K.: Natural variability of phytoplanktonic absorption in oceanic waters: Influence of the size structure of algal populations, Journal of Geophysical Research Oceans, 109, 45-50, 2004.

Chen, C., Gong, G., and Jan, S.: Seasonal and episodic effects (dust storm) on nutrient dynamics in the upwelling across shelf break on the East China Sea to the Kuroshio Water, northeastern Taiwan, AGU Spring Meeting, 2009,

Chen, Y. L. L.: Comparisons of primary productivity and phytoplankton size structure in the marginal regions of southern East China Sea, Continental Shelf Research, 20, 437-458, 2000.

Ciotti, A. M., Lewis, M. R., and Cullen, J. J.: Assessment of the relationships between dominant cell size in natural phytoplankton communities and the spectral shape of the absorption coefficient, Limnology and Oceanography, 47, 404-417, 2002.

Cleveland, J., and Weidemann, A. D.: Quantifying absorption by aquatic particles: A multiple scattering correction for glass‐fiber filters, Limnology and Oceanography, 38, 1321-1327, 1993.

Furuya, K., Kurita, K., and Odate, T.: Distribution of phytoplankton in the East China Sea in the winter of 1993, Journal of Oceanography, 52, 323-333, 1996.

Furuya, K., Hayashi, M., Yabushita, Y., and Ishikawa, A.: Phytoplankton dynamics in the East China Sea in spring and summer as revealed by HPLC-derived pigment signatures, Deep Sea Research Part II: Topical Studies in Oceanography, 50, 367-387, 2003.

Guidi, L., Stemmann, L., Jackson, G. A., Ibanez, F., Claustre, H., Legendre, L., Picheral, M., and Gorskya, G.: Effects of phytoplankton community on production, size, and export of large aggregates: A world‐ocean analysis, Limnology and Oceanography, 54, 1951-1963, 2009.

Hirata, T., Hardman-Mountford, N., Brewin, R., Aiken, J., Barlow, R., Suzuki, K., Isada, T., Howell, E., Hashioka, T., and

Noguchi-Aita, M.: Synoptic relationships between surface Chlorophyll-a and diagnostic pigments specific to phytoplankton functional types, Biogeosciences, 8, 311-327, 2011.

Hirawake, T., Takao, S., Horimoto, N., Ishimaru, T., Yamaguchi, Y., and Fukuchi, M.: A phytoplankton absorption-based primary productivity model for remote sensing in the Southern Ocean, Polar biology, 34, 291-302, 2011.

5 Hosmer Jr, D. W., Lemeshow, S., and Sturdivant, R. X.: Applied logistic regression, John Wiley & Sons, 2013.

Hu, C., Feng, L., and Lee, Z.: Uncertainties of SeaWiFS and MODIS remote sensing reflectance: Implications from clear water measurements, Remote sensing of environment, 133, 168-182, 2013.

IOCCG: Phytoplankton functional types from Space. In Sathyendranath, Shubha(Ed.), in, Reports of International Ocean-Colour Coordinating Group, no 15, 1-156, Dartmouth, Canada: IOCCG, 2014.

10 Jiang, Z., Liu, J., Chen, J., Chen, Q., Yan, X., Xuan, J., and Zeng, J.: Responses of summer phytoplankton community to drastic environmental changes in the Changjiang (Yangtze River) estuary during the past 50 years, Water Research, 54, 1-11, 2014.

Jiang, Z., Chen, J., Zhou, F., Shou, L., Chen, Q., Tao, B., Yan, X., and Wang, K.: Controlling factors of summer phytoplankton community in the Changjiang (Yangtze River) Estuary and adjacent East China Sea shelf, Continental Shelf Research, 101, 71-84, 2015.

15 Jiao, N., Yang, Y., Hong, N., Ma, Y., Harada, S., Koshikawa, H., and Watanabe, M.: Dynamics of autotrophic picoplankton and heterotrophic bacteria in the East China Sea, Continental Shelf Research, 25, 1265-1279, 2005.

Katoh, O., Morinaga, K., and Nakagawa, N.: Current distributions in the southern East China Sea in summer, Journal of Geophysical Research Atmospheres, 105, 8565-8574, 2000.

Kiørboe, T.: Turbulence, phytoplankton cell size, and the structure of pelagic food webs, Advances in marine biology, 29, 1-72, 1993.

Lamont, T., Brewin, R., and Barlow, R.: Seasonal variation in remotely-sensed phytoplankton size structure around southern Africa, Remote Sensing of Environment, 204, 617-631, 2018.

Lee, Z., Carder, K. L., and Arnone, R. A.: Deriving inherent optical properties from water color: a multiband quasi-analytical algorithm for optically deep waters, Applied optics, 41, 5755-5772, 2002.

25 Lee, Z., Lubac, B., Werdell, J., and Arnone, R.: An update of the quasi-analytical algorithm (QAA_v5), International Ocean Color Group Software Report, 1-9, 2009.

Lee, Z., Shang, S., Hu, C., and Zibordi, G.: Spectral interdependence of remote-sensing reflectance and its implications on the design of ocean color satellite sensors, Applied optics, 53, 3301-3310, 2014.

Li, Y., Li, D. J., Tang, J. L., Wang, Y. M., Liu, Z. G., Ding, P. X., and He, S. Q.: Phytoplankton Distribution and Variation in the Yangtze River Estuary and Its Adjacent Sea, Enviromental Science, 28, 719-729 (in Chinese), 2007.

Liu, X., Xiao, W., Landry, M. R., Chiang, K. P., Wang, L., and Huang, B.: Responses of Phytoplankton Communities to Environmental Variability in the East China Sea, Ecosystems, 19, 832-849, 2016.

Lou, X., and Hu, C.: Diurnal changes of a harmful algal bloom in the East China Sea: Observations from GOCI, Remote Sensing of Environment, 140, 562-572, 2014.

Lovelock, J.: The revenge of gaia: earth's climate crisis & the fate of humanity, Basic Books, 2007.

Luan, Q. S., Sun, J., Song, S. Q., Shen, Z. L., and Yu, Z. M.: Canonical correspondence analysis of summer phytoplankton community and its environment in the Yangtze River Estuary, China, Journal of Plant Ecology 31, 445-450 (in Chinese), 2007.

5   Meister, G.: Calibration and characterization adjustments to the MODIS ocean color band by the OBPG, Proc. MODIS Sci. Team Meeting, 2011, 1-23,

Mitchell, B. G.: Algorithms for determining the absorption coefficient for aquatic particulates using the quantitative filter technique, Orlando'90, 16-20 April, 1990, 137-148,

Montagnes, D. J., Berges, J. A., Harrison, P. J., and Taylor, F.: Estimating carbon, nitrogen, protein, and chlorophyll a from
10   volume in marine phytoplankton, Limnology and Oceanography, 39, 1044-1060, 1994.

Mueller, J. L., Fargion, G. S., Mcclain, C. R., Pegau, S., Zanefeld, J. R. V., Mitchell, B. G., Kahru, M., Wieland, J., and Stramska, M.: Ocean Optics Protocols for Satellite Ocean Color Sensor Validation, Revision 4, Volume IV: Inherent Optical Properties: Instruments, Characterizations, Field Measurements and Data Analysis Protocols, 2003.

Sheldon, R., and Parsons, T. R.: A practical manual on the use of the Coulter Counter in marine research, Coulter Electronics
15   Sales, 1967.

Sieburth, J. M., Smetacek, V., and Lenz, J.: Pelagic ecosystem structure: heterotrophic compartments of the plankton and their relationship to plankton size fractions, Limnology and oceanography, 23, 1256-1263, 1978.

Siswanto, E., Nakata, H., Matsuoka, Y., Tanaka, K., Kiyomoto, Y., Okamura, K., Zhu, J., and Ishizaka, J.: The long‐term freshening and nutrient increases in summer surface water in the northern East China Sea in relation to Changjiang
20   discharge variation, Journal of Geophysical Research Oceans, 113, 2601-2613, 2008.

Stæhr, P., Markager, S., and Sand-Jensen, K.: Pigment specific in vivo light absorption of phytoplankton from estuarine, coastal and oceanic waters, Marine Ecology Progress, 275, 115-128, 2004.

Su, J. L., and Yuan, Y. L.: Coastal Hydrology in China, Ocean Press, Beijing (in Chinese), 2005.

Sun, D., Hu, C., Qiu, Z., and Wang, S.: Reconstruction of hyperspectral reflectance for optically complex turbid inland lakes:
25   test of a new scheme and implications for inversion algorithms, Optics Express, 23, A718, 2015.

Sun, S., Lu, J., and Zhang, L.: Applications of flow cytometer in ecological studies of nano- and pico-phytoplankton, Chinese Journal of Ecology, 19, 72-78 (in Chinese), 2000.

Tang, J., Tian, G., Wang, X., Wang, X., and Song, Q.: The methods of water spectra measurement and analysis I: above-water method, Journal of Remote Sensing, 8, 37-44, 2004.

30   Tassan, S., and Ferrari, G. M.: An alternative approach to absorption measurements of aquatic particles retained on filters, Limnology and Oceanography, 40, 1358-1368, 1995.

Tassan, S., and Ferrari, G.: A sensitivity analysis of the 'Transmittance–Reflectance' method for measuring light absorption by aquatic particles, Journal of Plankton Research, 24, 757-774, 2002.

Taylor, A. H., and Joint, I.: A steady-state analysis of the 'microbial loop' in stratified systems, Marine Ecology Progress, 59,

1-17, 1990.

Tiwari, S. P., and Shanmugam, P.: An evaluation of models for the satellite-estimation of phytoplankton absorption coefficients in coastal/oceanic waters, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7, 364-371, 2014.

Uitz, J., Claustre, H., Morel, A., and Hooker, S. B.: Vertical distribution of phytoplankton communities in open ocean: An assessment based on surface chlorophyll, Journal of Geophysical Research: Oceans, 111, 2006.

Uitz, J., Huot, Y., Bruyant, F., Babin, M., and Claustre, H.: Relating phytoplankton photophysiological properties to community structure on large scales, Limnology and Oceanography, 53, 614, 2008.

Van Heukelem, L., and Hooker, S.: The importance of a quality assurance plan for method validation and minimizing uncertainties in the HPLC analysis of phytoplankton pigments, Phytoplankton Pigments, 195-256, 2011.

Van, H. L., and Thomas, C. S.: Computer-assisted high-performance liquid chromatography method development with applications to the isolation and analysis of phytoplankton pigments, Journal of Chromatography A, 910, 31-49, 2001.

Vidussi, F., Claustre, H., Manca, B. B., Luchetta, A., and Marty, J. C.: Phytoplankton pigment distribution in relation to upper thermocline circulation in the eastern Mediterranean Sea during winter, Journal of Geophysical Research: Oceans, 106, 19939-19956, 2001.

Wang, S., Ishizaka, J., Hirawake, T., Watanabe, Y., Zhu, Y., Hayashi, M., and Yoo, S.: Remote estimation of phytoplankton size fractions using the spectral shape of light absorption, Optics express, 23, 10301-10318, 2015.

Wang, S. Q., Ishizaka, J., Yamaguchi, H., Tripathy, S. C., Hayashi, M., Xu, Y. J., Mino, Y., Matsuno, T., Watanabe, Y., and Yoo, S. J.: Influence of the Changjiang River on the light absorption properties of phytoplankton from the East China Sea, Biogeosciences, 11, 1759-1773, 2014.

Yamaguchi, H., Kim, H. C., Son, Y. B., Sang, W. K., Okamura, K., Kiyomoto, Y., and Ishizaka, J.: Seasonal and summer interannual variations of SeaWiFS chlorophyll a in the Yellow Sea and East China Sea, Progress in Oceanography, 105, 22–29, 2012.

Zhang, H., Qiu, Z., Sun, D., Wang, S., and He, Y.: Seasonal and Interannual Variability of Satellite-Derived Chlorophyll-a (2000–2012) in the Bohai Sea, China, Remote Sensing, 9, 582, 2017.

Zhang, J., Liu, S., Ren, J., Wu, Y., and Zhang, G.: Nutrient gradients from the eutrophic Changjiang (Yangtze River) Estuary to the oligotrophic Kuroshio waters and re-evaluation of budgets for the East China Sea Shelf, Progress in Oceanography, 74, 449-478, 2007.

Tables

Table 1. Parameter β_i values for the PSC model development.

	N	R	RMSE	β_0	β_1	β_2	β_3	β_4
Micro	170	0.89	0.11	1.05	3.48	-4.34	-13.09	16.04
Nano	170	0.70	0.11	-	-	-	-	-
Pico	170	0.84	0.11	-2.56	-1.52	1.24	-25.87	-1.86

Table 2. Results of the matchup comparison of satellite $R_{rs}(\lambda)$ with *in situ* measurements.

	Wavelength (nm)									
	412	**443**	**469**	**488**	**531**	**547**	**555**	**645**	**667**	**678**
N	21	21	21	21	21	21	21	21	21	21
R	0.46	0.73	0.85	0.88	0.95	0.96	0.97	0.90	0.86	0.85
RMSE	0.0026	0.0019	0.0016	0.0016	0.0011	0.0011	0.0012	0.00081	0.00077	0.00079
MAPE(%)	47.33	36.90	27.25	19.92	16.39	14.90	18.41	54.86	91.39	111.3

Table 3. Comparison of original satellite R_{rs} and reconstructed R_{rs}^{rc} data with *in situ* values at 412 and 443 nm.

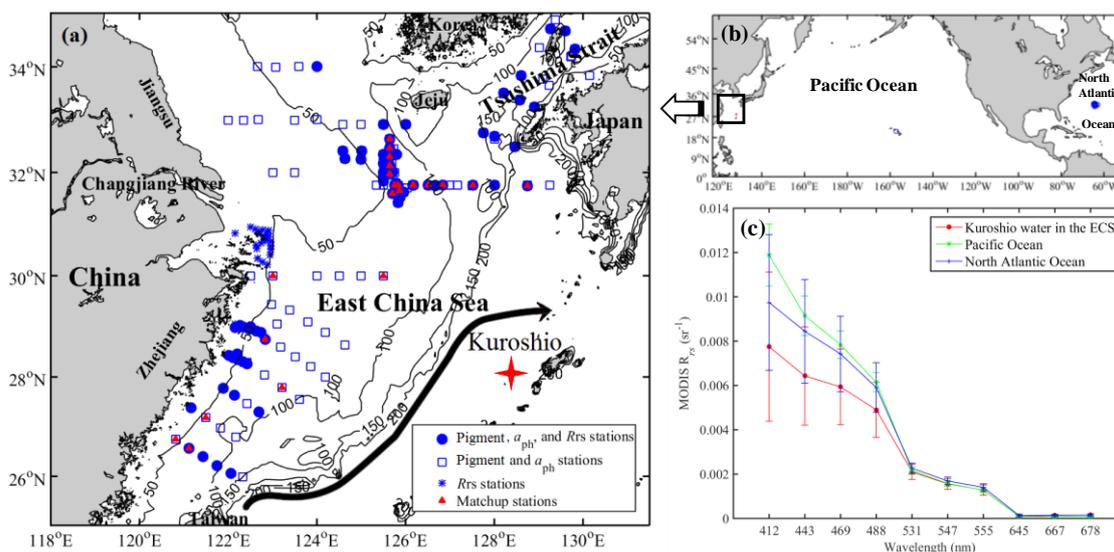
wavelength	N	original satellite R_{rs}			reconstructed satellite R_{rs}^{rc}		
		R	RMSE	MAPE(%)	R	RMSE	MAPE(%)
412 nm	21	0.46	0.0026	47.33	0.70	0.0019	35.15
443 nm	21	0.73	0.0019	36.90	0.80	0.0017	34.53

Table 4. Results of the matchup comparison for Fig. 6.

wavelength	N	satellite-derived a_{ph}^{rc}				satellite-derived a_{ph}			
		R	RMSE	MAPE(%)	PVP (%)	R	RMSE	MAPE(%)	PVP (%)
412 nm	22	0.80	0.27	15.35	100	0.58	0.56	58.15	72.73
443 nm	22	0.83	0.23	14.87	100	0.48	0.67	63.95	72.73
469 nm	22	0.85	0.21	11.62	100	0.62	0.55	48.35	81.82
488 nm	22	0.84	0.22	11.52	100	0.78	0.37	31.93	90.91
531 nm	22	0.80	0.31	20.82	86.36	0.86	0.73	45.57	54.55
547 nm	22	0.69	0.43	20.38	90.91	0.69	0.52	35.71	63.64

Table 5. Correlation coefficients between the monthly mean three phytoplankton size fractions and SST from 2003 to 3016 for the subareas. These relationships are statistically significant at the 95% confidence interval.

Region	*N*	Micro	Nano	Pico
MCJR	168	**-0.59**	**0.58**	**0.54**
MSR	168	**-0.76**	0.44	**0.65**
KR	168	-0.31	**-0.66**	**0.64**

Figures

5   **Fig. 1. Distribution of *in situ* and matchup dataset (a); locations of sampling stations collected in the North Pacific and North Atlantic oceans from the NASA SeaBASS archive (b); Mean $R_{rs}(\lambda)$ spectral of the 14-years (2003-2016) MODIS data in Kuroshio water (red star in a) and in the North Pacific and North Atlantic oceans (blue circles in b) (c). Error bars represent standard deviations of the means.**

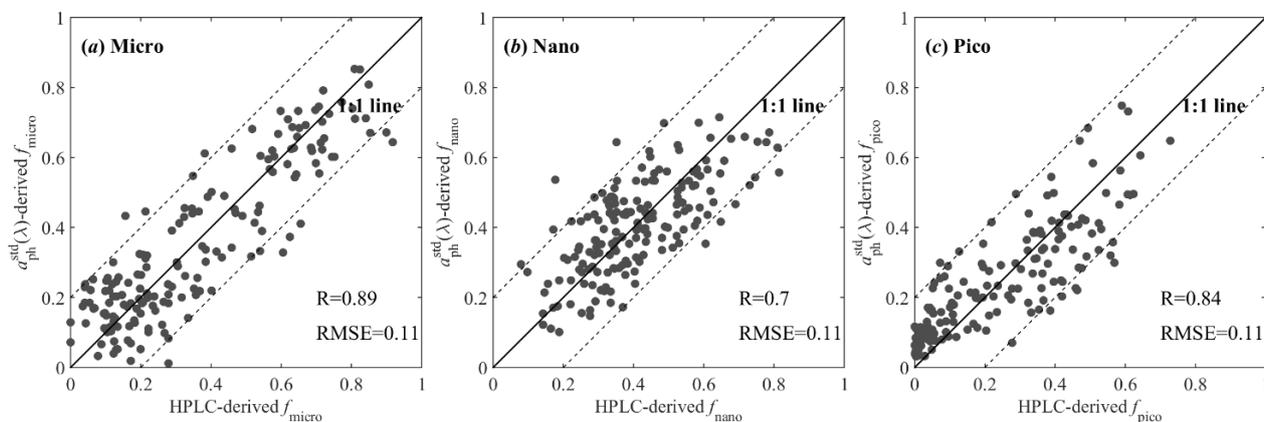
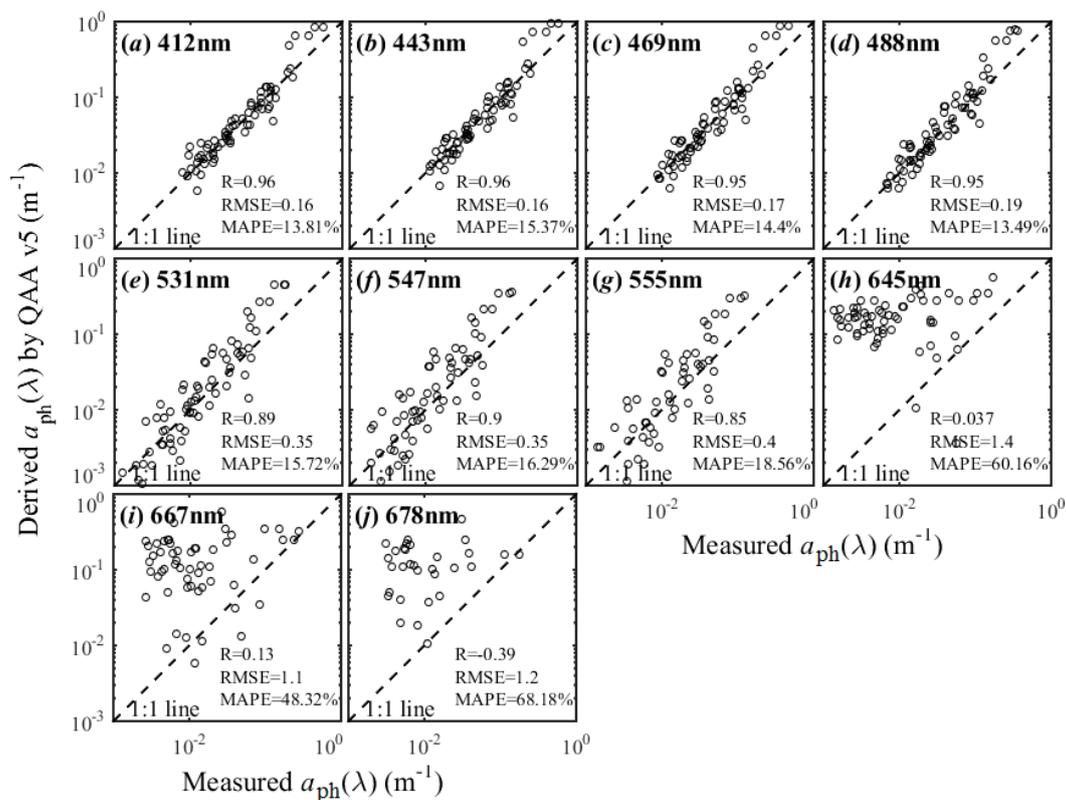
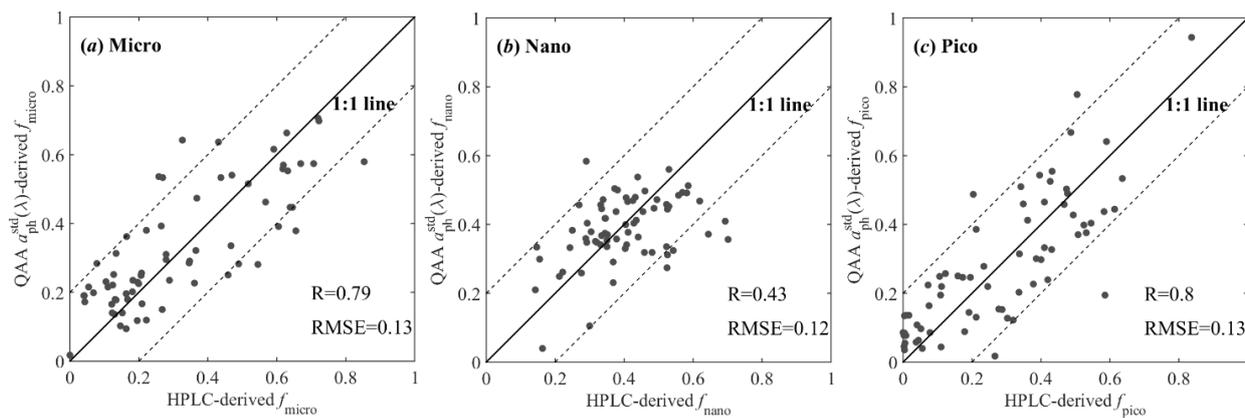
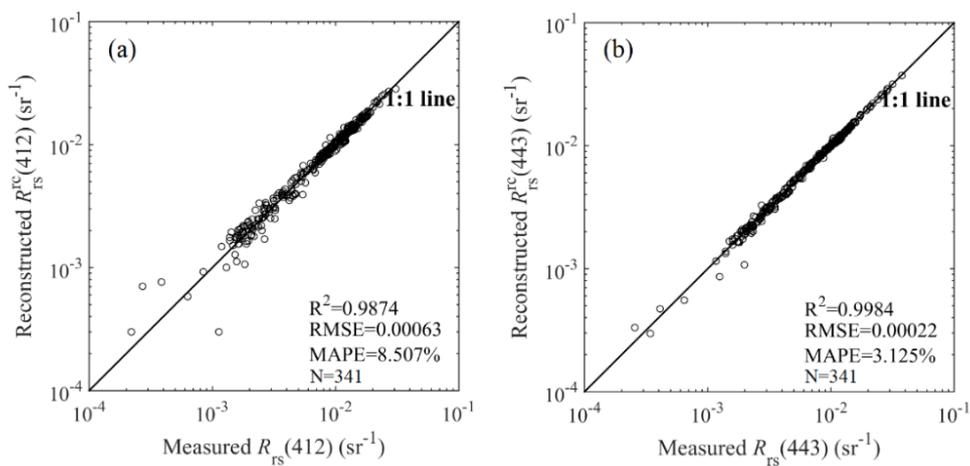
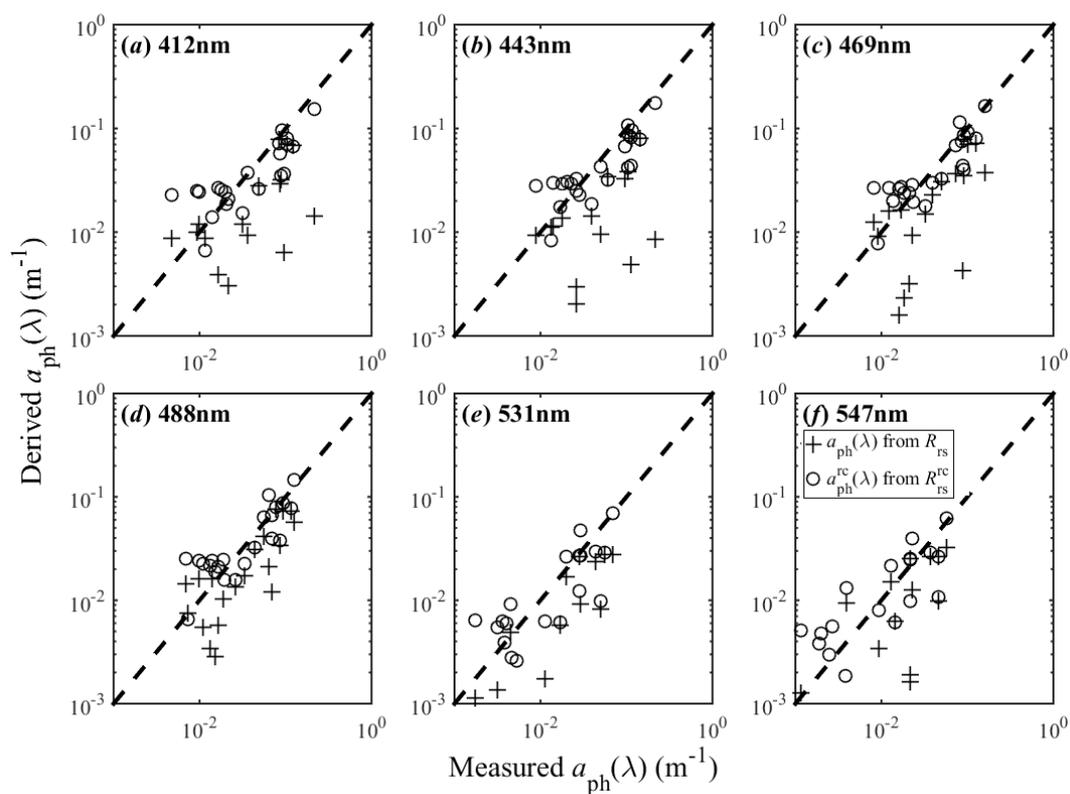



Fig. 2. Comparison between *in situ* $a_{ph}^{std}(\lambda)$-derived and HPLC-derived PSC for micro- (a), nano- (b), and pico-phytoplankton (c). Dashed lines represent the ± 20% fraction range relative to the 1:1 line.

5   **Fig. 3. Comparison of a_{ph} derived from R_{rs} using QAA_v5 with measured a_{ph} values.**

5   **Fig. 4. Comparisons of PSC modeled using a_{ph} derived from R_{rs} with HPLC-derived values for micro- (a), nano- (b),**

and pico-phytoplankton (c).

5   **Fig. 5. Reconstructed R_{rs}^{rc} versus measured R_{rs} data at 412 nm (a) and 443 nm (b).**

5   **Fig. 6. Comparison of the satellite-derived a_{ph} (crosses) and satellite-derived a_{ph}^{rc} (open circles) with measured values at**

412, 443, 469, 488, 531, and 547 nm.

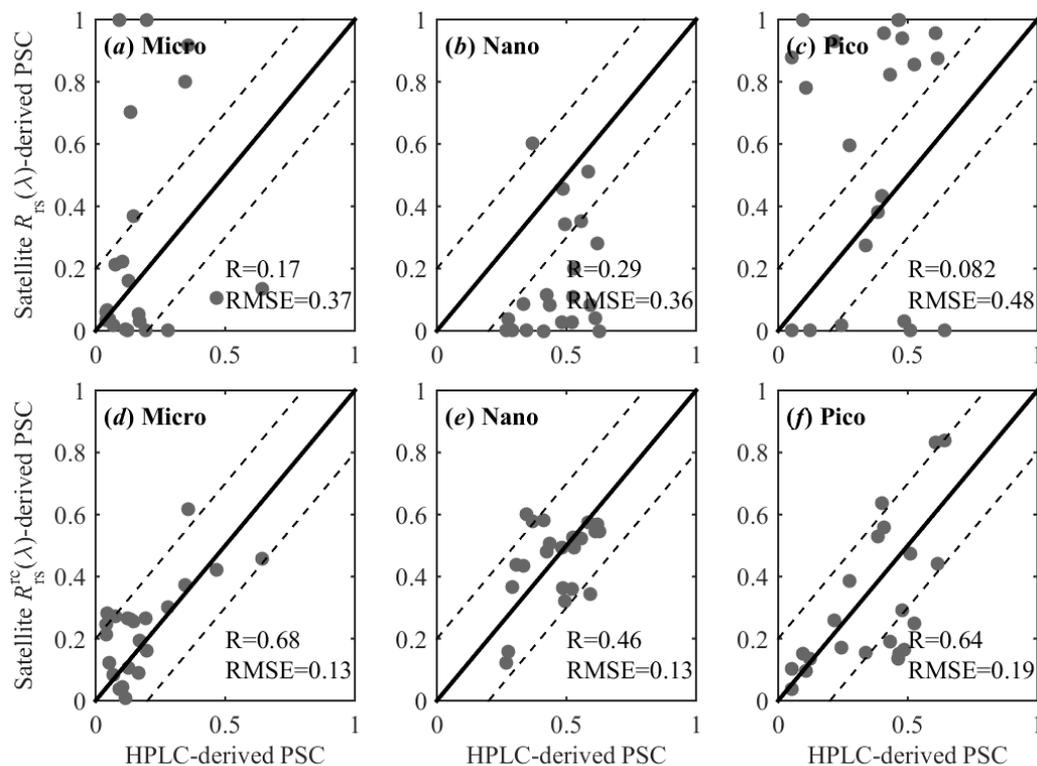



Fig. 7. Comparison of the satellite-derived PSC using original satellite R_{rs} data (a-c) and those using satellite R_{rs}^{rc} data (d-f)

5   with the HPLC-derived values.

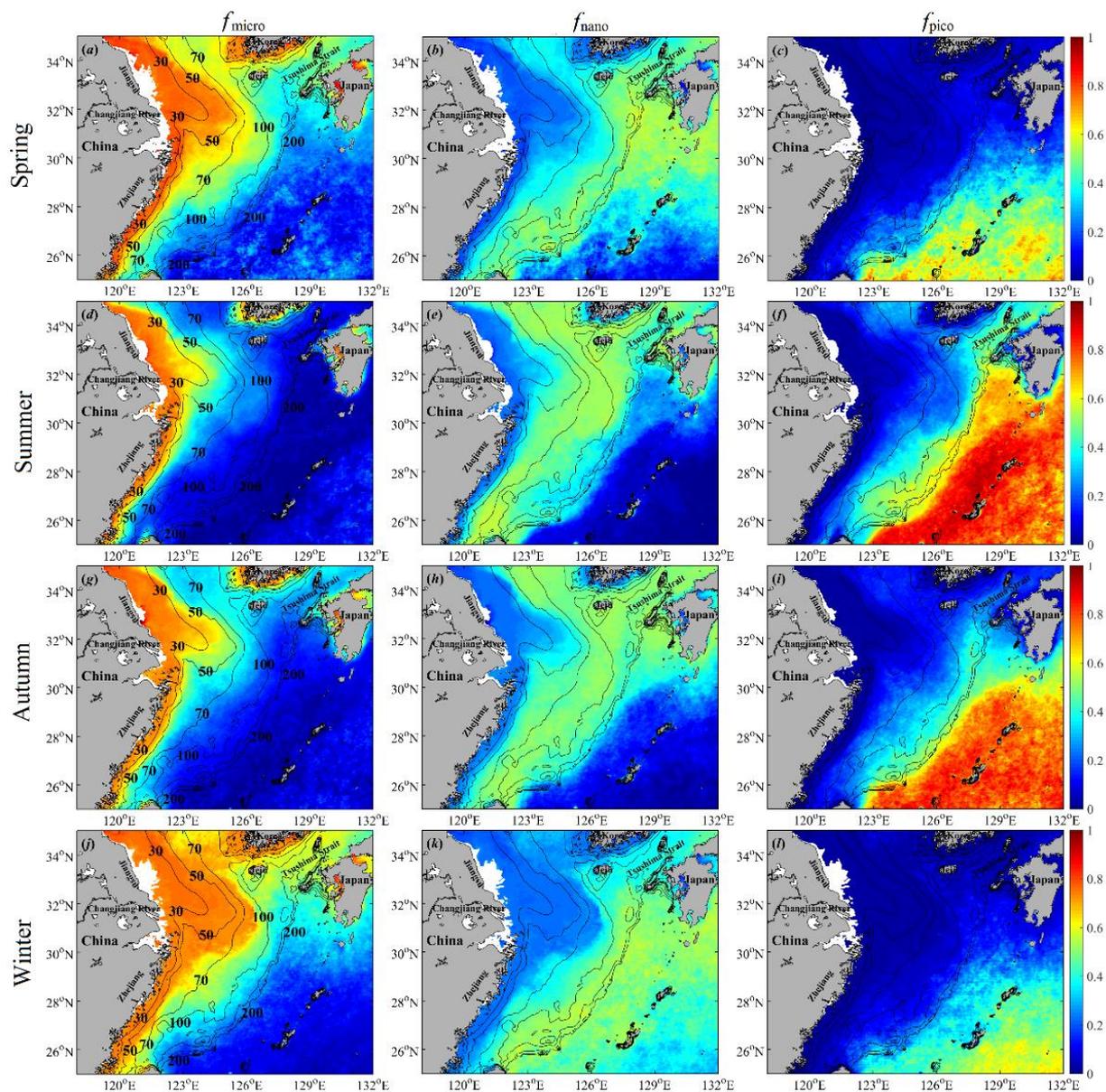



Fig. 8. Seasonal patterns of the PSC in the ECS during 2003-2016.

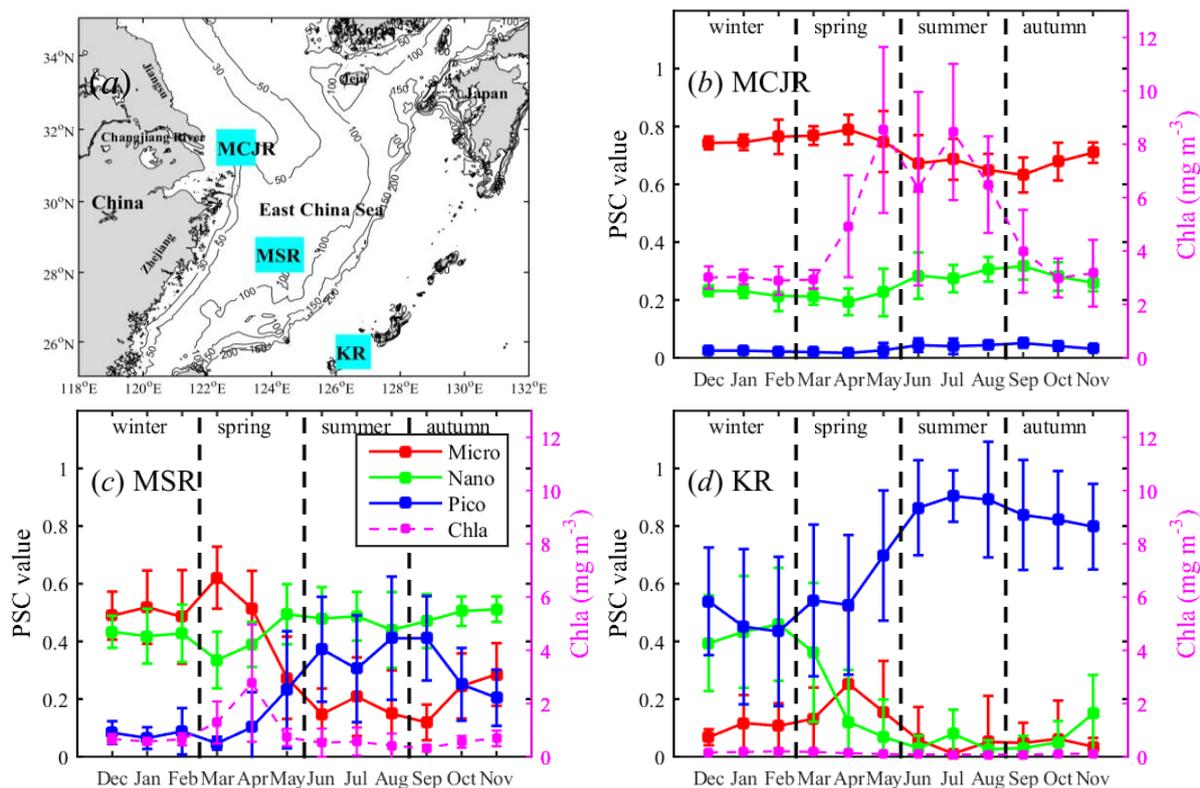



Fig. 9. The sampling location of subareas, which are marked by cyan color (a). Monthly climatological PSC(2003-2016)

in the mouth area of Changjiang river (MCJR) (b), middle shelf region (MSR) (c), and the Kuroshio region (KR) (d).

5 **Monthly climatological Chla concentration in the subareas are shown by magenta color. Error bars indicate standard**

deviations of the means.