

Temporal-spatial change of phytoplankton size structure from satellite observations in the Bohai and Yellow Seas of China during 2003–2024*

Yueqi WANG^{1,2}, Jiawen HAO^{1,2}, Yujue WANG³, Dongyan LIU^{3,**}

¹ Yantai Institute of Coastal Zone Research, Chinese Academy of Sciences, Yantai 264003, China

² University of Chinese Academy of Sciences, Beijing 100190, China

³ State Key Laboratory of Estuarine and Coastal Research, Institute of Eco-Chongming, East China Normal University, Shanghai 200241, China

Received Nov. 18, 2025; accepted in principle Dec. 26, 2025; accepted for publication Jan. 16, 2026

© Chinese Society for Oceanology and Limnology, Science Press and Springer-Verlag GmbH Germany, part of Springer Nature 2026

Abstract Phytoplankton size structure is a key ecological indicator for assessing marine ecosystems status and their response to environmental change; however, long-term monitoring of phytoplankton size has been constrained by limitations in observational data. To address this gap, this study developed a 22-year (2003–2024) daily satellite-derived phytoplankton size class (PSC) dataset for the Bohai and Yellow seas of China (BYS) using an in-situ optimized abundance-based model. Key findings are: (1) micro-phytoplankton dominated the coastal waters all-year round, whereas offshore water exhibited seasonal alternation between pico-phytoplankton (summer) and micro-phytoplankton (other seasons); (2) a significant long-term phytoplankton size miniaturization was detected, with the contributions of micro- and nano-phytoplankton declined at a rates of -0.031% per year (non-significant, $P>0.05$) and -0.049% per year (significant, $P<0.01$), respectively, while pico-phytoplankton increased by 0.055% per year (significant, $P<0.01$); (3) through explainable machine learning model, sea surface temperature was identified as the dominant driver of PSC variations, exhibiting highly nonlinear relationships and distinct threshold effects with shifts in phytoplankton size structure. This validated, long-term satellite PSC product provides novel insights into the environmental forcing of phytoplankton community restructuring in the BYS and offering a critical baseline for assessing ecosystem responses to climate change and anthropogenic pressure.

Keyword: phytoplankton size class (PSC); remote sensing; model optimization; long-term change; Bohai and Yellow seas

1 INTRODUCTION

Marine phytoplankton, responsible for approximately half of global total primary production, forms the fundamental basis of marine food webs and biogeochemical cycles (Field et al., 1998). The increasing pressures from climate change and anthropogenic activities are profoundly altering marine ecosystems, with phytoplankton communities being particularly vulnerable to these environmental perturbations (Hoegh-Guldberg and Bruno, 2010). Characterizing phytoplankton dynamic has therefore become imperative for

projecting future ecosystem trajectories under global change scenarios (Barnett et al., 2001; Smetacek and Cloern, 2008). Size-based classification offers a powerful framework for investigating phytoplankton community structure and its functional linkages with physical processes and biogeochemical cycles (Uitz

* Supported by the National Natural Science Foundation of China (No. 42030402), the Natural Science Foundation of Shandong Province (No. ZR2022MD028), and the Open Research Fund of Key Laboratory of Land and Sea Ecological Governance and Systematic Regulation, Ministry of Ecology and Environment (No. 202501)

** Corresponding author: dyliu@sklec.ecnu.edu.cn

et al., 2008; IOCCG, 2014). Physiological differences are pronounced across size classes: smaller cells benefit from high surface area-to-volume ratios that confer growth efficiency, while larger cells exhibit greater photosynthetic capacity under resource-replete conditions (e.g., rich in light and nutrient) but experience higher sinking losses (Marañón et al., 2012). Following the operational classification of Sieburth et al. (1978), phytoplankton is commonly divided into three size classes: pico- ($<2\ \mu\text{m}$), nano- ($2\text{--}20\ \mu\text{m}$), and micro-phytoplankton ($>20\ \mu\text{m}$). The phytoplankton size classes (PSC) differentially mediate processes such as new production and particulate organic carbon export fluxes (Guidi et al., 2009). However, significant knowledge gaps remain in our understanding of their spatiotemporal distributions and environment responses, largely due to observational challenges of obtaining synoptic, long-term PSC dataset in ocean.

Traditional PSC quantification relies on multiple in-situ methodologies, including microscopy, flow-cytometry, size-fractionated filtration, and high-performance liquid chromatography (HPLC), each with distinct analytical trade-off (IOCCG, 2014). Although precise, these laboratory techniques are operationally constrained: requiring specialized equipment, lengthy processing, and substantial financial and labor investments. These limitations inherently restrict the spatiotemporal coverage of in-situ PSC observations, creating significant gaps in our understanding of basin-scale phytoplankton community dynamics (Turner et al., 2021). In contrast, satellite remote sensing provides the potential for synoptic and continuous spatiotemporal coverage. Consequently, deriving PSC information from satellite ocean color data has become an active research focus. Several PSC algorithms have been developed for both global (Brewin et al., 2011, 2015; Sun et al., 2023) and regional applications (Sun et al., 2019a; Liu et al., 2021; Turner et al., 2021).

Algorithms for retrieving PSC from ocean color remote sensing generally fall into three categories: spectral-based, ecological-based, and abundance-based approaches (Brewin et al., 2010; Devred et al., 2011; Sun et al., 2023). These algorithms differ in their theoretical frameworks and input requirement (IOCCG, 2014). Abundance-based PSC algorithms rely on empirical or semi-empirical relationships between PSC composition and total chlorophyll-*a* (Chl-*a*) concentration, typically

derived from in-situ observations (Brewin et al., 2011; Hirata et al., 2011). Given that Chl *a* is widely available and validated satellite product, abundance-based methods offer a straightforward and practical approach to estimating PSC from remote sensing data, often yielding reliable results (Turner et al., 2021; Sun et al., 2023; de Lima Oliveira et al., 2025). Regional studies have successfully re-parameterized abundance-based PSC models calibrated with local in-situ data in diverse seas including the Arabian Sea (Sahay et al., 2017), the Red Sea (Gittings et al., 2019), the Atlantic Ocean (Moore and Brown, 2020), and the continental shelf seas of China (Sun et al., 2017, 2019b; Liu et al., 2021). These studies collectively confirm the broad applicability of abundance-based PSC models across diverse ocean regions. However, observed variations in optimal model parameters between regions underscore the necessity of local calibration prior to PSC dynamic analysis.

The Bohai and Yellow seas (BYS) are shallow, semi-enclosed continental shelf seas in northern China (depth $<100\ \text{m}$) (Fig.1). It was commonly partitioned into coastal and offshore provinces using the 50-m isobaths, a recognized hydrological and ecological boundary that approximates the seaward limit of strong coastal influence, separating the well-mixed, turbid, and nutrient-rich shallow waters from the more stratified offshore basin. This division aligns with standard zoning practices for the Yellow Sea Large Marine Ecosystem and is supported by studies that document significant shifts in biological and physicochemical properties across this boundary (Wang and Gao, 2020). As typical marginal seas, the BYS experience synergistic pressures from global warming, terrestrial inputs, atmospheric deposition, and ocean dynamics, which modulating water column stability, light availability, and nutrient regimes, leading to pronounced phytoplankton variability. Consequently, phytoplankton dynamics in the BYS have been extensively studied via field surveys and satellite observations (Fu et al., 2009; Hao et al., 2019; Jiang et al., 2019; Sun et al., 2019b; Wang et al., 2021; Zhai et al., 2021). Research has delineated nearshore-offshore biomass gradients driven primarily by terrestrial nutrients (Liu and Wang, 2013), identified divergent seasonal control mechanisms (Yamaguchi et al., 2012; Wang and Gao, 2020), and revealed a significant non-monotonic long-term Chl-*a* trend with a regime shift from a positive to a negative around 2011, attributed to the changing nutrient states (Wang et al., 2021;

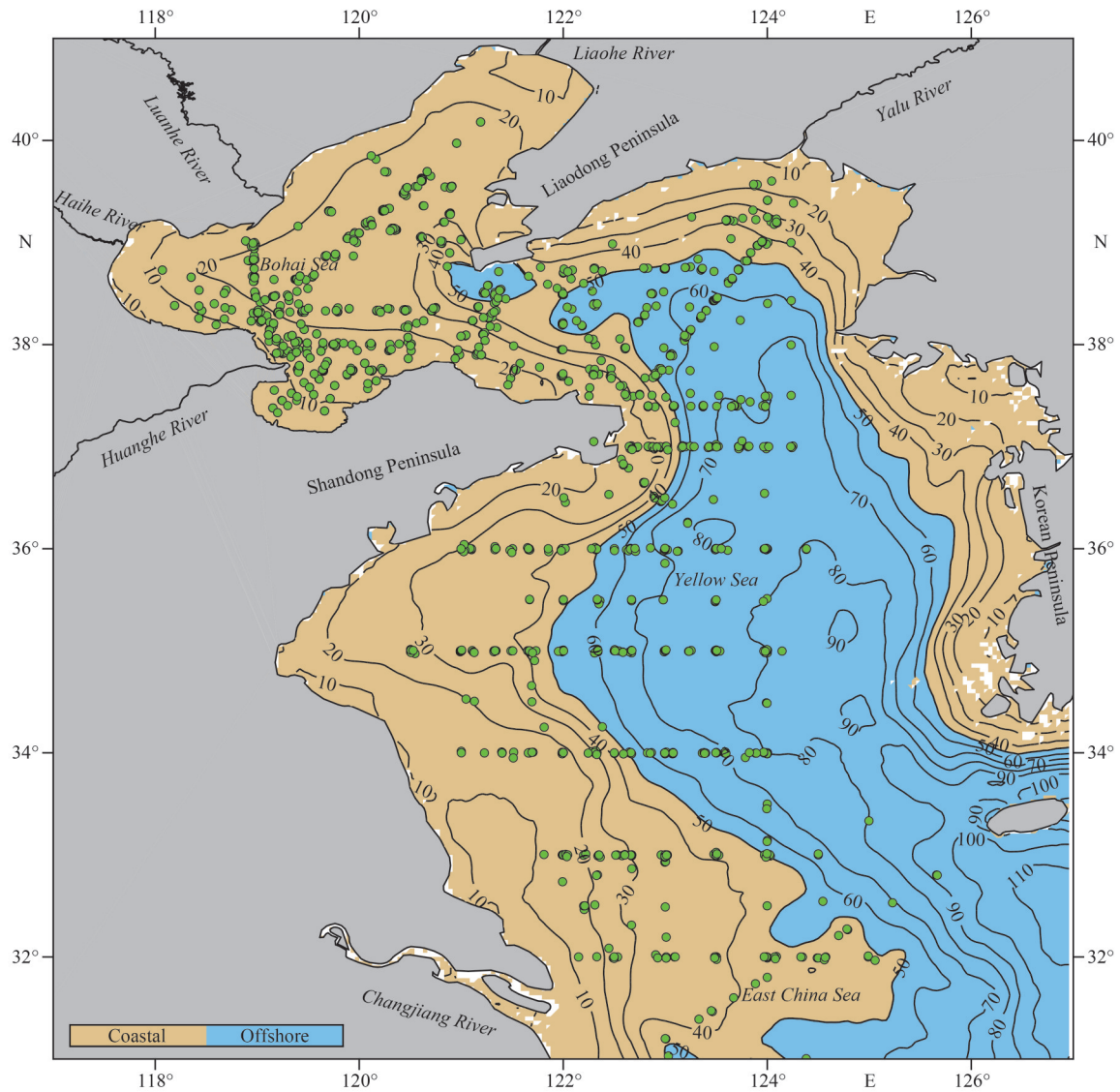


Fig.1 Study area and bathymetry of the Bohai and Yellow seas (BYS)

The 50-m isobaths (solid black line) delineates the coastal (<50 m) and offshore (>50 m) provinces. Green dots denote the locations of in-situ pigment sampling stations. Bathymetric data were from the General Bathymetric Chart of the Oceans (GEBCO): <https://www.gebco.net/>.

Zhai et al., 2021).

In contrast to phytoplankton biomass dynamics in BYS, research on PSC variability remains comparatively limited, particularly concerning long-term trends and their driving mechanisms. Recent advances in ocean color remote sensing and the accumulation of in-situ PSC observations have facilitated the development of regional satellite retrieval algorithms (Sun et al., 2017, 2019b), yielding preliminarily insights into regional PSC spatiotemporal patterns (Sun et al., 2019a, b; Liu et al., 2021). Spatially, eutrophic coastal waters exhibit high total Chl *a* dominated by large-sized PSC, whereas oligotrophic offshore waters favor small-

sized PSC. Temporally, micro-phytoplankton prevails during cold seasons, while nano- and pico-phytoplankton dominate warmer periods. These patterns align with size-dependent resource competition under varying environmental conditions, reflecting physiological and ecological trade-offs (Agawin et al., 2000; Marañón et al., 2015). Although satellite-derived synoptic PSC patterns were mostly consistent with previous field surveys (Fu et al., 2009; Jiang et al., 2019), several uncertainties persist. For instance, while regime shifts in total biomass have been documented over the last two decades (Wang et al., 2021; Zhai et al., 2023), some studies reported no significant long-

term PSC trend (Sun et al., 2017). Furthermore, PSC model parameters showed substantial variations across studies, likely stemming from methodological differences, including model calibration limited to warm-season data (April–October) (Sun et al., 2017, 2018, 2019b), or insufficient tailoring of models to BYS-specific conditions (Liu et al., 2021).

In summary, optimizing an abundance-based PSC model with an extensive and updated in-situ dataset from the BYS is essential for generating more accurate remote-sensing PSC products and paving the way for further investigations into the long-term variability of PSC. Leveraging these enhanced satellite products, our study aims to provide a comprehensive analysis of PSC variability and its potential drivers in the BYS. Specifically, we compile a large in-situ dataset encompassing total Chl-*a* concentrations and PSC fractions, coupled with concurrent sea surface temperature (SST) observations. Using these updated data, we refine and recalibrate an existing SST-dependent PSC model (Wang et al., 2024) for application in the BYS. The objectives of this research are threefold: (1) to establish a reliable long-term PSC dataset for the BYS; (2) to quantify decade-scale PSC variability; and (3) to identify environmental drivers on PSC changes.

2 MATERIAL AND METHOD

Figure 2 provides a schematic overview of the study workflow. Briefly, datasets comprise in-situ surveys and satellite remote sensing observations. Key methodologies include the optimization of SST-dependent abundance-based PSC model and spatiotemporal analysis of satellite-derived PSC products. Detailed descriptions of in-situ samples, satellite-derived variables and analytical methodologies are provided in the subsequent subsections.

2.1 In-situ data collection and processing

Field measurements were acquired through 23 research cruises conducted in the BYS between 2013 and 2023, yielding 1 611 discrete water samples. Measured parameters included SST obtained using conductivity-temperature-depth (CTD) equipment and phytoplankton pigment quantified via HPLC. Details of the HPLC pigment measurements have been documented in many previous studies (Liu et al., 2021; Wang et al.,

2024). Briefly, water samples were filtered and preserved at -80°C during the cruises, and later analyzed in the laboratory using an Agilent 1260 Infinity II HPLC system following the protocols described by Zapata et al. (2000) and Furuya et al. (2003). Analyses quantified total Chl-*a* concentration (sum of monovinyl Chl *a*, divinyl Chl *a*, and chlorophyllide-*a*) and seven diagnostic pigments (includes fucoxanthin, peridinin, 19'-hex-fucoxanthin, 19'-but-fucoxanthin, alloxanthin, chl *b*, and zeaxanthin).

Strict quality control was implemented: (1) Chl-*a* concentrations lower than 0.001 mg/m^3 were set to 0.001 mg/m^3 ; (2) accessory pigments concentrations lower than 0.0001 mg/m^3 were set to 0.0001 mg/m^3 ; (3) samples were retained only if the difference between the Chl-*a* concentration and the total concentration of accessory pigments was less than 30% of the total pigment concentration (Aiken et al., 2009). These criteria excluded unrealistic low concentrations data and samples with anomalous pigment covariance (Trees et al., 2000; Brewin et al., 2015). The final curated dataset comprised 1 406 samples spanning a range of months (including January, March, April, May, June, July, August, September, November, and December), with broad spatial coverage across the BYS (Figs. 1 & 2a). The dataset includes a broad range of Chl-*a* and SST values, with Chl-*a* concentrations ranging from 0.03 to 18.34 mg/m^3 and SST ranging from 0.01 to 30.21°C .

2.2 Satellite-derived dataset and processing

2.2.1 Satellite-derived Chl-*a* product

Sea surface Chl-*a* concentrations were derived using a regional generalized additive model (GAM) algorithm applied to MODIS daily remote sensing reflectance (R_{rs}) data (Wang et al., 2017, 2019). This GAM-based Chl-*a* algorithm implements a nonlinear smoothed regression framework, with in-situ Chl-*a* measurements as the response variable and visible R_{rs} bands (412, 443, 469, 488, 531, 547, 555, 645, 667, and 678 nm) plus water depth as predictors (Wang et al., 2017). Comparative validation demonstrated superior performance over alternative Chl-*a* products for the BYS, achieving a root mean square error (RMSE) of 0.21 mg/m^3 and a coefficient of determination (R^2) of 0.81 (Wang et al., 2019). The algorithm's accuracy and robustness makes it particularly suitable for analyzing Chl-*a*

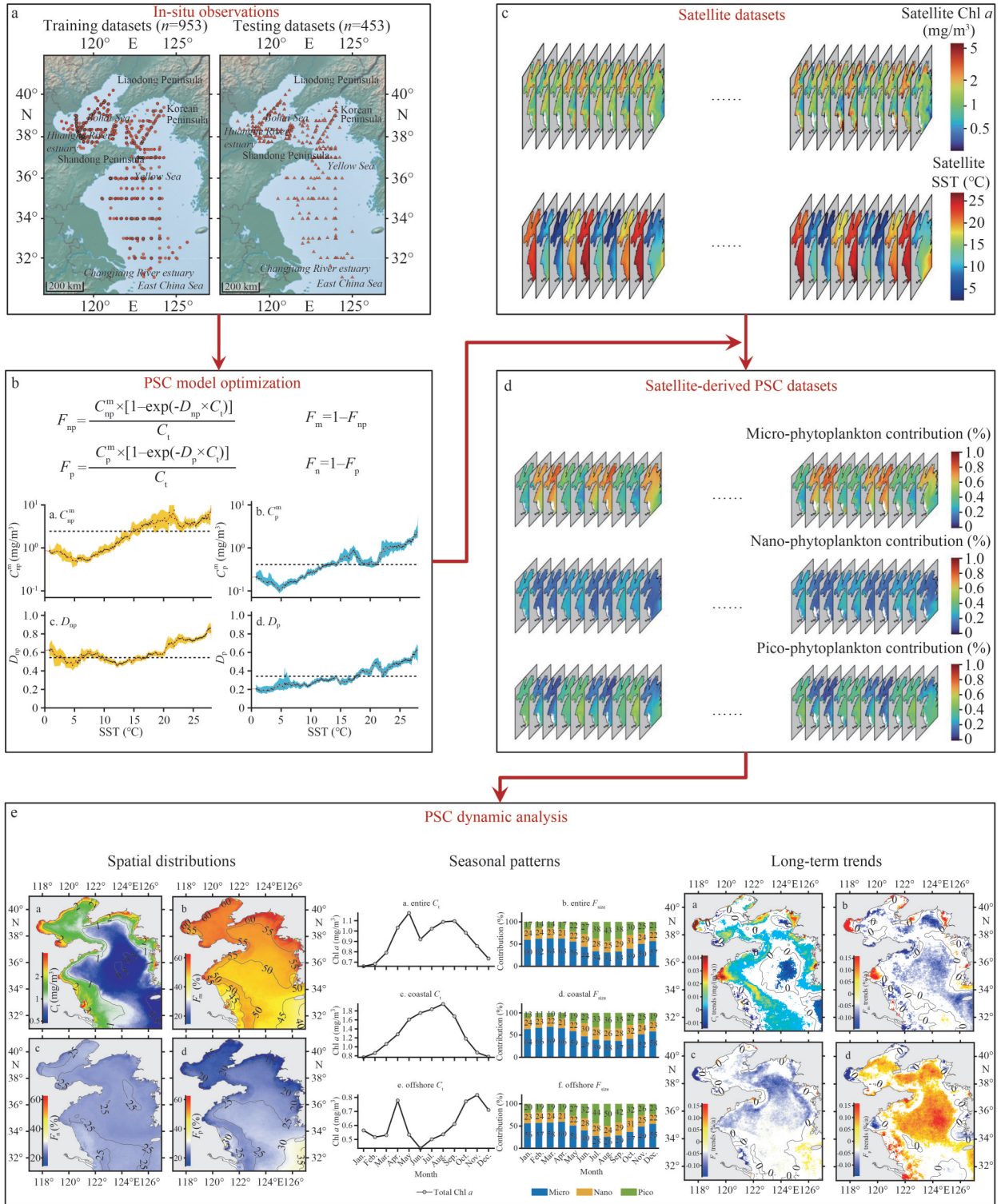


Fig.2 Schematic diagram illustrating the workflow of this study

a. in-situ measurements were partitioned into training and testing datasets (n : number of samples); b. the abundance-based phytoplankton size class (PSC) model was optimized; c. satellite-derived Chl a and SST data were input; d. a long-term satellite-derived PSC dataset was generated; e. spatiotemporal dynamics of PSC were analyzed. Chl a : chlorophyll a ; SST: sea surface temperature; C_t : total Chl- a concentration; C_m : micro-phytoplankton Chl- a concentration, C_n : nano-phytoplankton Chl- a concentration; C_p : pico-phytoplankton Chl- a concentration; F_m : micro-phytoplankton contribution; F_n : nano-phytoplankton contribution; F_p : pico-phytoplankton contribution. C_{size} : size-fractionated Chl- a concentrations; F_{size} : size-fractionated contributions.

dynamics in the BYS (Wang et al., 2021, 2023b; Lv et al., 2022; Lin et al., 2023). For the present study,

daily MODIS Chl- a products with a spatial resolution of $1/24^\circ \times 1/24^\circ$ (approximately $4 \text{ km} \times$

4 km) spanning from 1 January 2003 to 31 December 2024 were processed and used for further analysis.

2.2.2 Satellite-derived sea surface temperature (SST)

The satellite-derived SST data used in this study were sourced from NOAA's Optimum Interpolation (OI) daily SST dataset (Reynolds and Smith, 1994; Reynolds et al., 2007). These data were freely accessible via the NOAA's National Climatic Data Center website (<https://www.ncei.noaa.gov/data/>). For this study, we utilized the SST product composed exclusively of advanced very high resolution radiometer (AVHRR) infra-red satellite measurements spanning from 1 January 2003 to 31 December 2024. The SST dataset was resampled to a grid resolution of approximately $1/24^\circ \times 1/24^\circ$ using nearest-neighbor interpolation to align with the satellite-derived PSC dataset.

2.2.3 Satellite-derived photosynthetically active radiation (PAR)

Daily photosynthetically active radiation (PAR) data derived from MODIS/Aqua were acquired from the NASA OceanColor website (<http://oceancolor.gsfc.nasa.gov/>) for the period from 1 January 2003 to 31 December 2024. These PAR products were generated using algorithms developed by Frouin et al. (2003). Monthly PAR composites were then created and resampled to a $1/24^\circ \times 1/24^\circ$ spatial grid to align with the satellite-derived Chl-*a* dataset.

2.2.4 Satellite-derived total suspended matter (TSM)

Daily total suspended matter (TSM) concentration data were calculated from the daily MODIS Rrs dataset using an empirical algorithm developed by Zhang et al. (2010). The monthly TSM images were composited from the daily images with the spatial resolution of $1/24^\circ \times 1/24^\circ$ to align with the satellite-derived Chl-*a* dataset.

2.2.5 Mixed layer depth (MLD)

Mixed layer depth (MLD) dataset were acquired from the ARMOR3D level-4 reanalysis products distributed by the Copernicus Marine Environment Monitoring Service (CMEMS) (<https://resources.marine.copernicus.eu>). These data are available on a $1/4^\circ \times 1/4^\circ$ spatial resolution and were estimated using a density threshold equivalent to a 0.2°C temperature decrease relative to 10-m reference depth (de Boyer Montégut et al., 2004). For consistency with other datasets, the monthly MLD products were resampled to a $1/24^\circ \times 1/24^\circ$ spatial resolution (approximately $4\text{ km} \times 4\text{ km}$).

2.2.6 Sea surface wind (SSW)

Sea surface wind (SSW) data were sourced from the Cross-Calibrated Multi-Platform (CCMP) version 3.0 gridded surface 10-m wind vector dataset (<https://www.remss.com/measurements/ccmp/>). This dataset comprises 6-hourly sea surface vector wind maps, gridded at a resolution of $1/4^\circ \times 1/4^\circ$, and are synthesized from various microwave scatterometers and radiometers (including QuikSCAT, ASCAT, SSM/I, AMSR, SSMIS, and WindSat) via variational data assimilation (Atlas et al., 2011). The monthly SSW data were composited from the 6-hourly images and were resampled to a $1/24^\circ \times 1/24^\circ$ grid to align with the satellite-derived Chl-*a* dataset.

2.3 In-situ PSC estimation and preprocessing

To support model re-parameterization and validation, in-situ PSC were derived from quality-controlled pigment measurements using the Diagnostic Pigment Analysis (DPA) method (Uitz et al., 2008; Brewin et al., 2010). Within the DPA framework, 1 406 quality-controlled HPLC pigment samples utilized as input data. A multivariate linear regression was employed, with total Chl-*a* concentration serving as the dependent variable and seven diagnostic pigments as independent variables. This process yielded weighting coefficients of 1.14, 1.21, 0.52, 0.01, 4.53, 1.57, and 1.43 for the pigments of fucoxanthin, peridinin, 19'-hex-fucoxanthin, 19'-but-fucoxanthin, alloxanthin, Chl *b*, and zeaxanthin, respectively. Total Chl *a* was subsequently partitioned into three PSCs (pico-plankton, nano-plankton, and micro-plankton) by applying these coefficients to calculate the weighted contribution of each pigment to the respective size classes. This procedure generated a curated, quality-assured in-situ PSC dataset comprising 1 406 samples, suitable for subsequent model calibration and validation. The parameters of the in-situ PSC dataset include the total Chl-*a* concentration (C_t), size-fractionated Chl-*a* concentrations (C_{size}), and size-fractionated contributions (F_{size}). Specifically, C_t comprises micro-phytoplankton concentration (C_m), nano-phytoplankton concentration (C_n), and pico-phytoplankton concentration (C_p); while F_{size} consists of the contributions of micro- (F_m), nano- (F_n), and pico-phytoplankton (F_p).

To enable comparison with satellite data, the 1 406 in-situ PSC samples were matched daily with remotely sensed SST and Chl-*a* data (detailed in subsequent sections) based on temporal coincidence

(same day) and the closest spatial coordinates. To ensure the quality of the satellite retrievals, in-situ measurements were matched to satellite imagery using 3×3 pixel window, and only match-ups containing five or more valid pixels within the window were retained. Furthermore, the coefficient of variation (CV) for satellite-derived Chl *a* across the valid pixels was calculated, and match-ups with a Chl-*a* CV greater than 0.15 were excluded. Based on these criteria, 453 satellite in-situ match-ups were selected to form a testing dataset for validating the remotely sensed PSC estimates (Fig.2a). The remaining 953 in-situ PSC samples were utilized to re-calibrate the SST-dependent abundance-based PSC model (Wang et al., 2024), constituting the training dataset (Fig.2a).

2.4 PSC model optimization and satellite application

The SST-dependent, abundance-based PSC model, initially proposed by Brewin et al. (2017) and refined specific for BYS by Wang et al. (2024), was re-parameterized using a training dataset comprising 953 in-situ PSC samples (Fig.2b). The performance of the re-parameterized model was evaluated using the testing dataset comprising 453 in-situ and satellite matchups, as illustrated in Fig.2a. Upon adjustment of the model parameters, the optimized PSC model was applied to satellite-derived daily Chl *a* and concurrent SST images, thereby generating remotely-sensed daily PSC datasets of the BYS. As a result, monthly PSC time series were derived by calculating the median of all daily PSC images for each month.

2.5 Mann-Kendall test for trend detection

The non-parametric seasonal Mann-Kendall (MK) test was employed to detect linear trends in the monthly data series (Hirsch et al., 1982; Hirsch and Slack, 1984). This approach was selected due to its robustness in handling the characteristically skewed distribution of Chl-*a* values and accounting for strong serial autocorrelation present in the monthly time series (Campbell, 1995; Saulquin et al., 2013). The magnitudes of trends were computed using Sen's method (Sen, 1968; Gilbert, 1987), which provides a non-parametric estimate of trend magnitude. As a well-established statistical approach, the MK test has been extensively applied in environmental studies to assess trends in various geographical variables, including the sea surface Chl *a*, in previous studies (Vantrepotte and Mélin,

2011; Pitarch et al., 2016). The methodology involves calculating MK statistics for each individual month, enabling the detection of persistent monotonic changes in the time series. The overall trend magnitude is determined by computing the median of monthly slope estimates. The overall trends are conventionally expressed as absolute changes per year. To enable comparison across varying concentration levels, we calculated relative linear trends (unit: %/a) by normalizing the Sen's slope against the median value of the fitted line.

2.6 Random forest (RF) regression analysis

RF regression, an ensemble machine learning algorithm developed by Breiman (2001), operates by constructing multiple decision trees on randomly selected subsets of both samples and features from the dataset, with final predictions obtained through averaging. This method has demonstrated particular efficacy in capturing complex nonlinear relationships between variables (Gao et al., 2023), and have been widely applied in numerous ecological studies, including phytoplankton biomass prediction and analysis of algal growth-environmental factors interactions (Xu et al., 2024). In the present study, we employed RF regression to: (1) quantify the relative importance of environmental drivers on PSC variations and (2) characterize nonlinear responses of PSC to environmental changes.

For our analysis, we processed monthly time series data of PSC (C_t and F_{size}) along with five key environmental factors (SST, PAR, TSM, MLD and SSW). These data were spatially averaged for coastal and offshore regions, respectively, yielding a final dataset of 269 matched samples. In RF, PSC parameters served as dependent variables while environmental factors acted as predictors. Recognizing the critical role of hyperparameter selection in data-driven modeling (Feurer et al., 2019), we implemented a Bayesian optimization approach to systematically tune the model parameters. This method, rooted in Bayes' theorem, provides an efficient and theoretically sound strategy for global optimization problems (Victoria and Maragatham, 2021), ensuring optimal model performance while maintaining computational efficiency.

2.7 SHapley Additive exPlanations (SHAP) analysis

SHAP provides a unified framework for both

local and global model interpretations, rooted in cooperative game theory's Shapley values (Lundberg and Lee, 2017). It enables quantification of each feature's contribution to model predictions by evaluating its marginal impact across all possible feature combinations. The SHAP value for a feature is calculated by comparing its value against a baseline value, typically the model's average prediction for that feature across all samples. By analyzing these values, we can understand the influence of each feature on the model's overall behavior. The magnitude of the SHAP value indicates the strength of the feature's impact on the model output: larger absolute values denote greater importance of that input feature. A positive SHAP value signifies that the feature increases the prediction, while a negative SHAP value signifies that it decreases the prediction. This approach facilitates transparent interpretation of the importance and complex relationships between environmental drivers and PSC in our model.

2.8 Statistical performance metrics

To assess the accuracy of satellite-derived total Chl-*a* and PSC products, three performance metrics were employed: the coefficient of determination (R^2), RMSE, and median absolute percentage error (MAPE). For the C_t , C_m , C_n , and C_p , R^2 and RMSE were computed in log-transformed space owing to their log-normal distributions (Seegers et al., 2018), whereas all the other metrics were calculated in linear space. The median absolute percentage error was chosen over mean absolute percentage error because the median offers a more robust and representative measure of central tendency in the presence of potential outliers.

3 RESULT

3.1 Re-parameterized PSC model and performance evaluation

The PSC model was re-parameterized using a training dataset comprising 953 in-situ samples, with the results illustrated in Fig.3. The SST-dependent, abundance-based PSC model was constructed by successively applying the abundance-based PSC model (Fig.2b) to the PSC data within each 0.1-°C SST bin over the range of 0.5–29.5 °C, resulted in 290 fit profiles (Fig.3a). In Fig.3a, the parameters C_{size} and F_{size} are plotted against total Chl *a* (C_t) for different SST ranges, confirming the clear influence

of SST on model parameters, aligning with findings from previous studies (Sun et al., 2019b; Liu et al., 2021; Wang et al., 2024). Lower SSTs were generally corresponded to a higher proportion of large-size phytoplankton (F_m), while higher SSTs favored smaller size classes (F_n and F_p). The relationship between SST and PSC parameters was not strictly linear, as indicated by the overlapping fitted PSC curves across different SSTs (Fig.3a). Overall, effect of SST on PSC fractions was discernible across the full Chl-*a* range but become more pronounced at lower Chl-*a* levels.

Independent validation was conducted by comparing satellite-derived and in-situ measured PSC using 453 testing samples (Fig.3b). The results demonstrate that satellite-derived size-specific Chl-*a* concentrations (C_m , C_n , and C_p) and fractions (F_m , F_n , and F_p) agree well with in-situ observations. Some larger discrepancies persisted in the quantification of nano-phytoplankton, likely due to its spectral or diagnostic pigment overlap with the other two size classes. Nevertheless, the satellite products demonstrate reliable accuracy in capturing PSC variability and are suitable for resolving the spatiotemporal patterns of phytoplankton size structure in the BYS.

3.2 Spatial distribution of PSC

Figure 4 presents climatological spatial distribution (2003–2024) of PSC in the BYS, derived from climatological composite satellite data. The total Chl-*a* concentration (C_t) demonstrates a distinct cross-shelf gradient, decreasing progressively from nutrient-rich coastal waters to oligotrophic offshore provinces (Fig.4a). Micro-phytoplankton dominates the PSC composition (51.4% basin-wide average), accounting for 40%–60% of total biomass in most areas (Fig.4b). Nano- and pico-phytoplankton exhibited comparable contributions, each representing 20%–30% of total biomass (Fig.4c & d). Notably, Phytoplankton size composition displayed obvious spatial heterogeneity, with larger cell sizes predominating in the Bohai Sea and northern Yellow Sea, while the southern Yellow Sea showed reduced micro-phytoplankton fraction and increased dominance of pico-phytoplankton (Fig.4). In contrast, the nano-phytoplankton (F_n) maintains stable proportions (25%) throughout the BYS. These satellite-derived PSC spatial distributions show good agreement with findings from previous in-situ studies (Fu et al., 2009; Jiang et al., 2019),

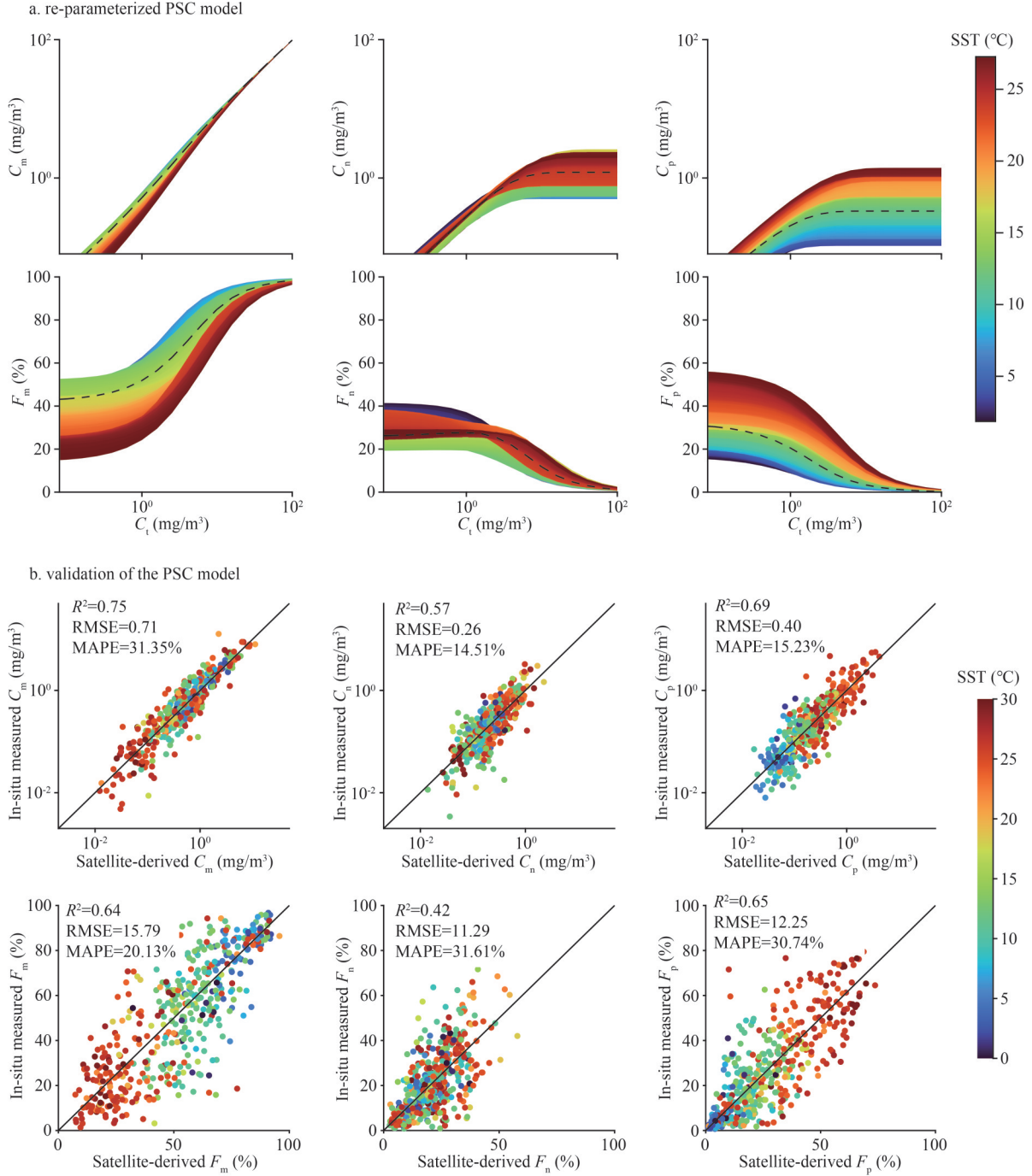


Fig.3 Re-parameterized PSC model parameters as a function of SST (a); validation of satellite-derived size-fractionated Chl-*a* concentrations (C_{size}) and contributions (F_{size}) against in-situ measurements (b) (performance metrics and 1:1 line are shown)

Chl *a*: chlorophyll *a*; SST: sea surface temperature; C_t : total Chl-*a* concentration; C_m : micro-phytoplankton Chl-*a* concentration, C_n : nano-phytoplankton Chl-*a* concentration; C_p : pico-phytoplankton Chl-*a* concentration; F_m : micro-phytoplankton contribution; F_n : nano-phytoplankton contribution; F_p : nano-phytoplankton contribution.

validating the remote sensing approach for PSC analysis.

3.3 Seasonal pattern of PSC

The seasonal variability of PSC distributions in

the BYS was synoptically delineated using climatological monthly satellite composites (2003–2024), as shown in Fig.5. The spatial distributions of C_t exhibited a persistent nearshore-offshore gradient year-round, with gradient intensity peaking during

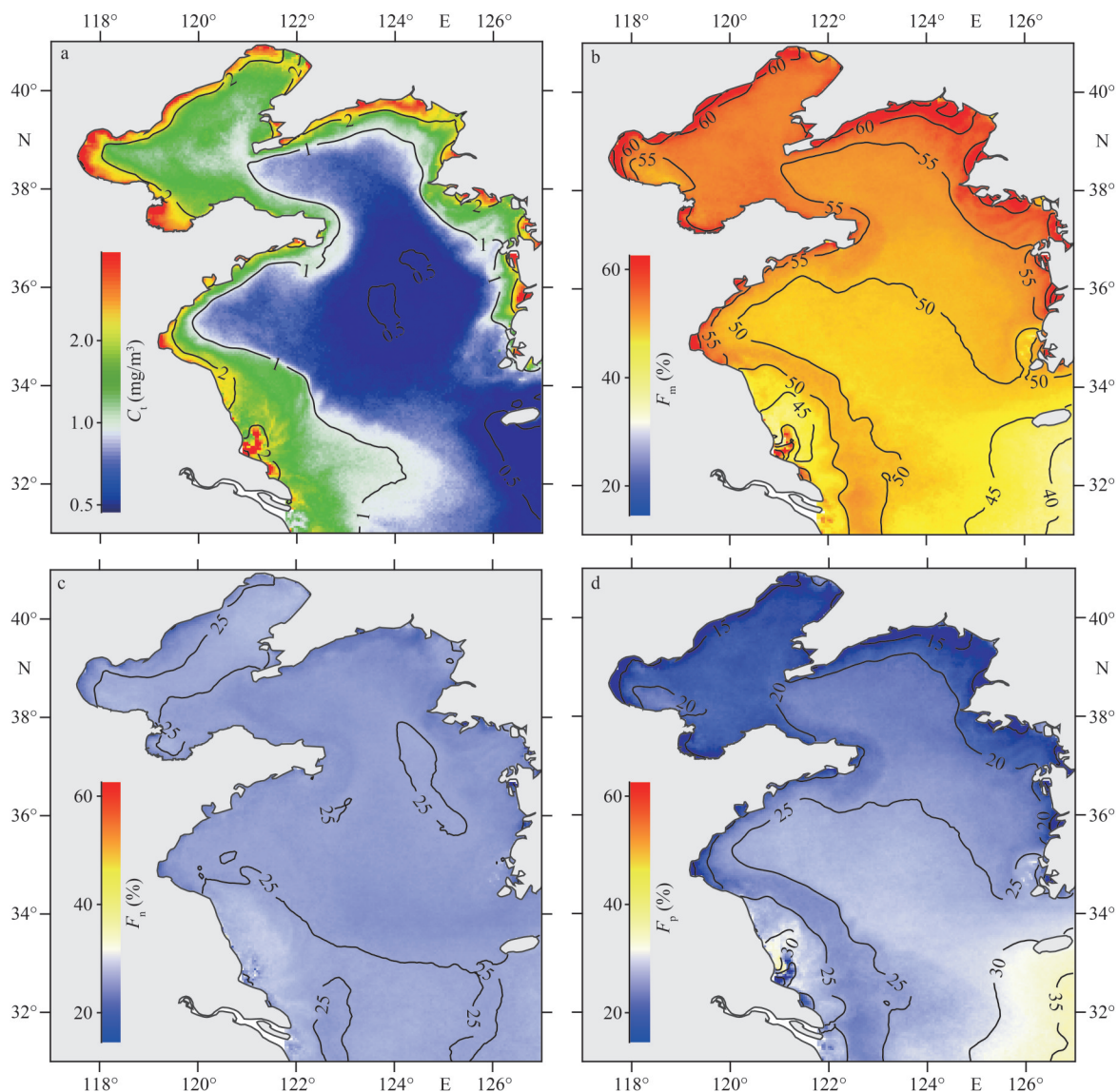


Fig.4 Climatological mean spatial distributions of total Chl *a* and fractional contributions

a. total Chl-*a* concentration; b. micro-phytoplankton contribution; c. nano-phytoplankton contribution; d. pico-phytoplankton contribution.

summer months (June–August) and diminishing in winter (December–February). Micro-phytoplankton dominates from October to June, with its contribution (F_m) generally exceeding 50%. In contrast, pico-phytoplankton predominate in offshore waters during the warm season (July–September), with its contribution (F_p) averaging approximately 40%, and nano-phytoplankton maintains intermediate contributions (20%–30%) across all seasons.

Figure 6 presents the climatological monthly variations of C_t and PSC means for the entire BYS, and its coastal and offshore sub-regions. The basin-wide means reveals a characteristic bimodal seasonal pattern in C_t (Fig.6a), featuring distinct

spring and autumn blooms separated by summer and winter troughs, consistent with Chl-*a* seasonal pattern reported in previous study (Wang and Gao, 2020). The offshore region similarly exhibited a bimodal C_t pattern (Fig.6i), with primary and secondary peaks occurring in spring and autumn respectively, separated by summer and winter minima. Notably, the summer trough of offshore C_t was more pronounced and prolonged compared to the entire C_t . The coastal C_t displayed a clear unimodal pattern characterized by a single warm-season maximum and cold-season minimum (Fig.6e), differing markedly from the bimodal patterns observed in both entire and offshore C_t . All the F_{size} showed comparable unimodal patterns, with

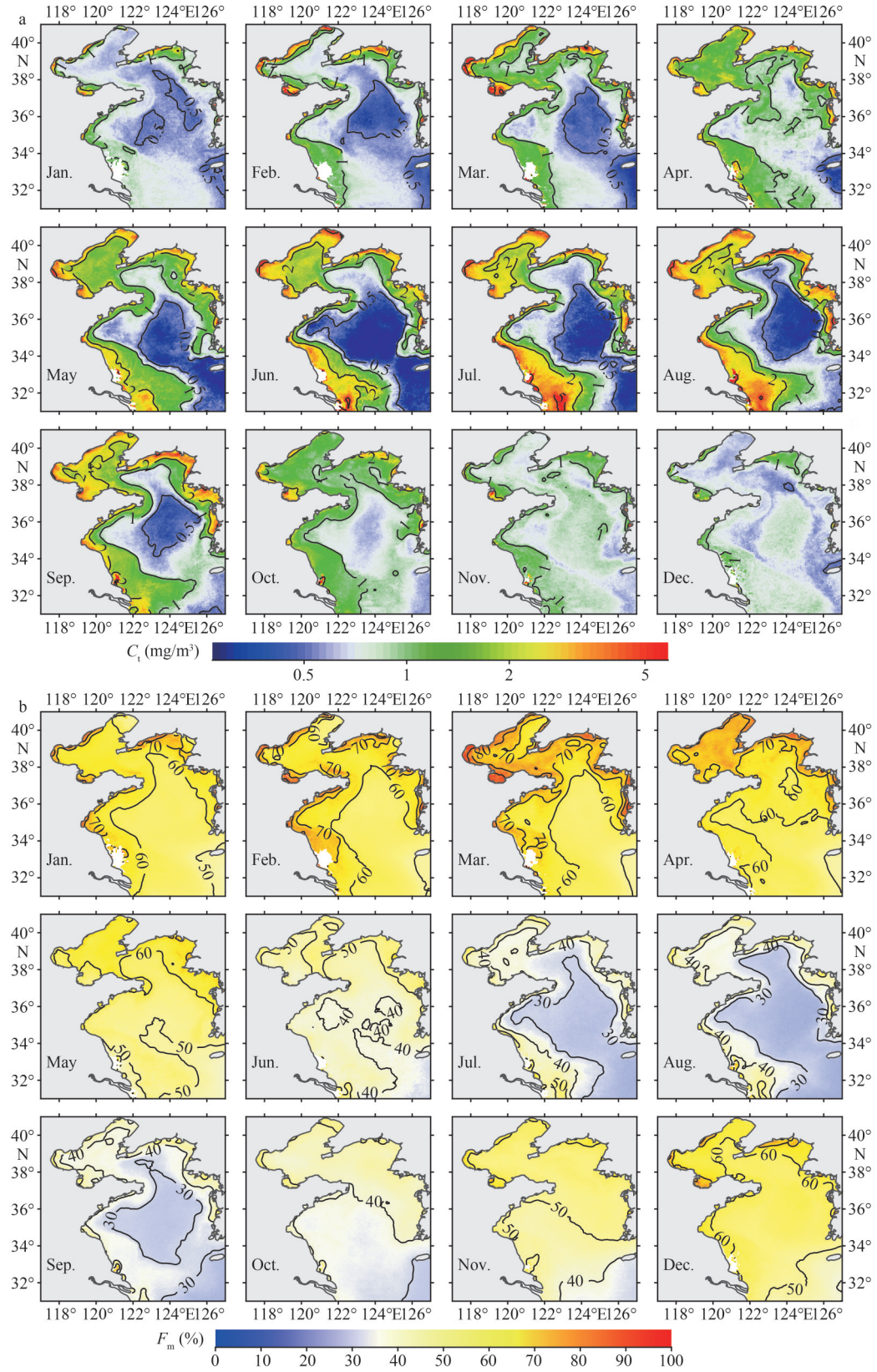
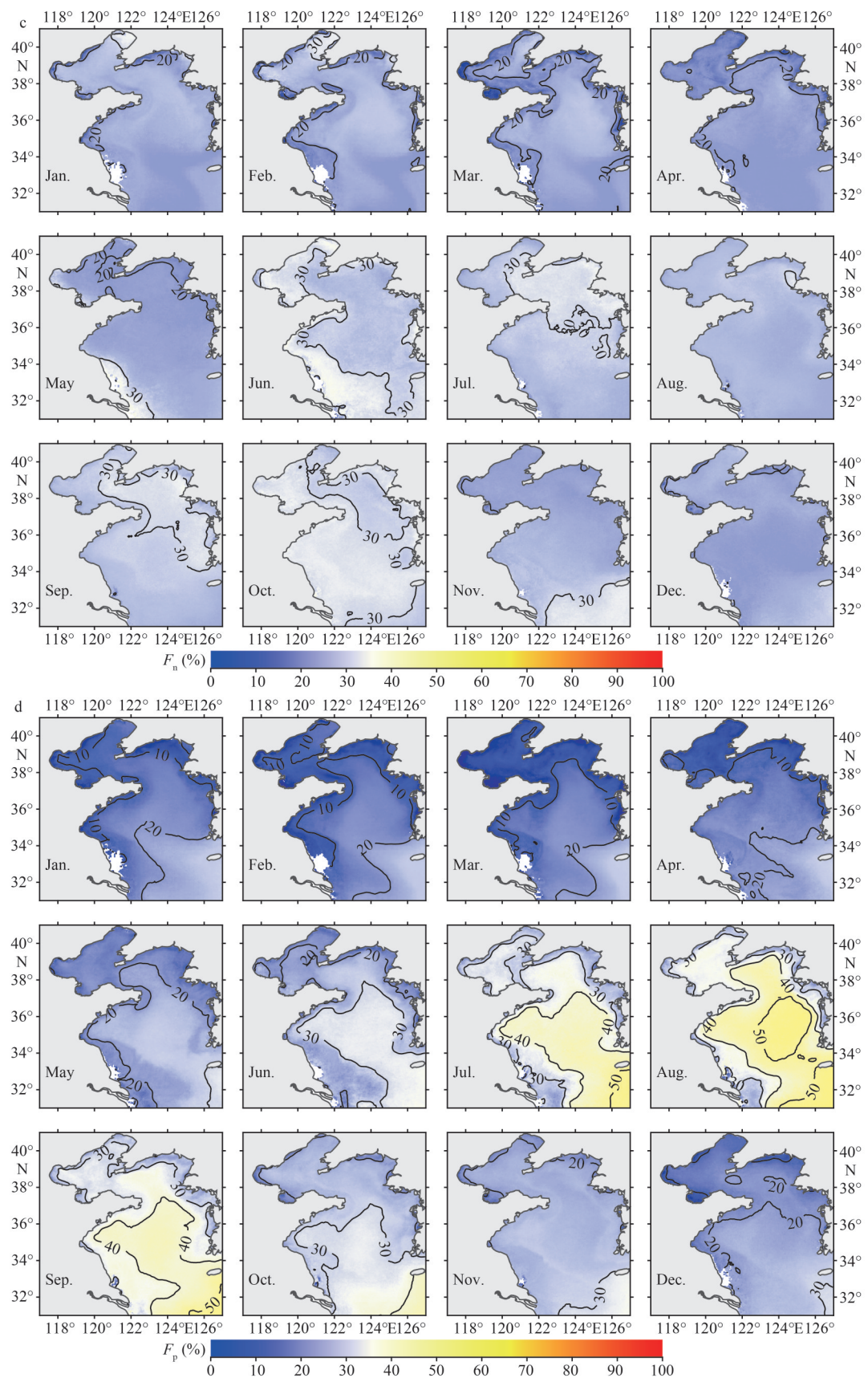


Fig.5 Climatological monthly composites (2003–2024) of total Chl-*a* concentration (C) (a), and fractional contributions of micro-phytoplankton (F_m) (b), nano-phytoplankton (F_n) (c), and pico-phytoplankton (F_p) (d) in the BYS; the spatially dominant phytoplankton size class for each month (e)

To be continued

Fig.5 Continued



To be continued

Fig.5 Continued

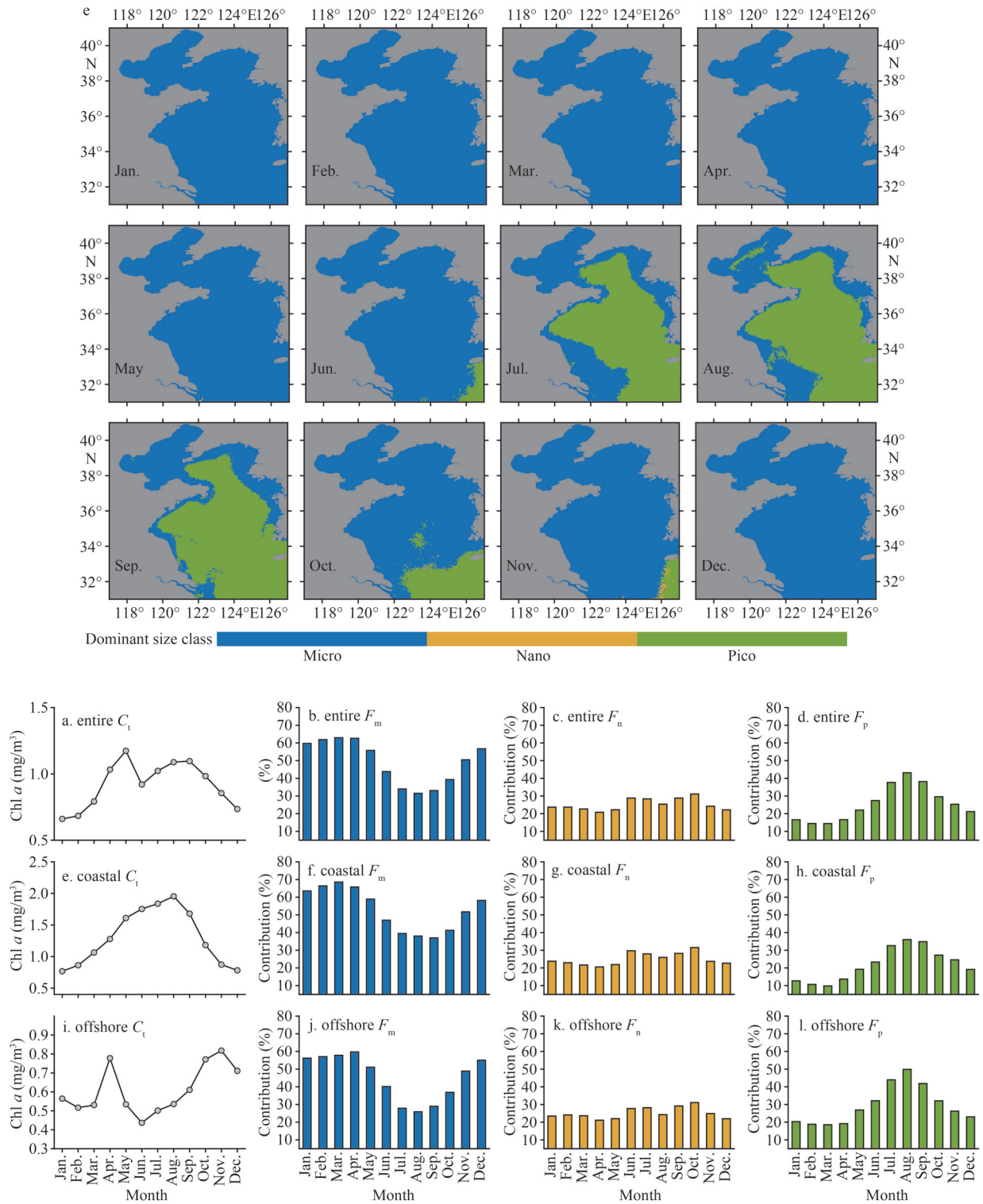


Fig.6 Climatological monthly means of total Chl-*a* concentration and contributions of each size across the entire BYS and its two sub-regions, as defined in Fig.1

C_t : total Chl-*a* concentration; F_m : micro-phytoplankton contribution; F_n : nano-phytoplankton contribution; F_p : pico-phytoplankton contribution.

larger phytoplankton predominating during winter-spring and smaller size classes dominating in

summer-autumn. A striking contrast emerged between coastal and offshore F_{size} : micro-

phytoplankton maintained year-round dominance in coastal region, whereas offshore waters exhibited seasonal shifts between pico-phytoplankton dominance (July–September) and micro-phytoplankton dominance (remaining months). These characteristics underscore the ecological predominance of micro-phytoplankton in coastal environments, where they consistently contribute higher biomass relative to other phytoplankton size classes. The satellite-derived PSC seasonal patterns align with findings in previous studies (Fu et al., 2009; Jiang et al., 2019).

3.4 Long-term trend in PSC

The MK test was applied to the monthly PSC

data series spanning 2003–2024 to detect and quantify linear trends, and the spatial distributions of trend slopes were shown in Fig.7. For C_t (Fig.7a), 45.95% of the area presented significant linear trends, in which significant positive and negative trends occupied 38.72% and 7.13% of the basin area, respectively. The significant positive C_t trends mainly prevail in coastal waters, where some nearshore patches demonstrated particularly strong positive trends. The significant negative C_t trends spatially concentrated in several offshore areas including central northern yellow sea, central southern yellow sea, and southeast open waters (Fig.7a). For F_{size} (Fig.7b–d), there are 36.77%,

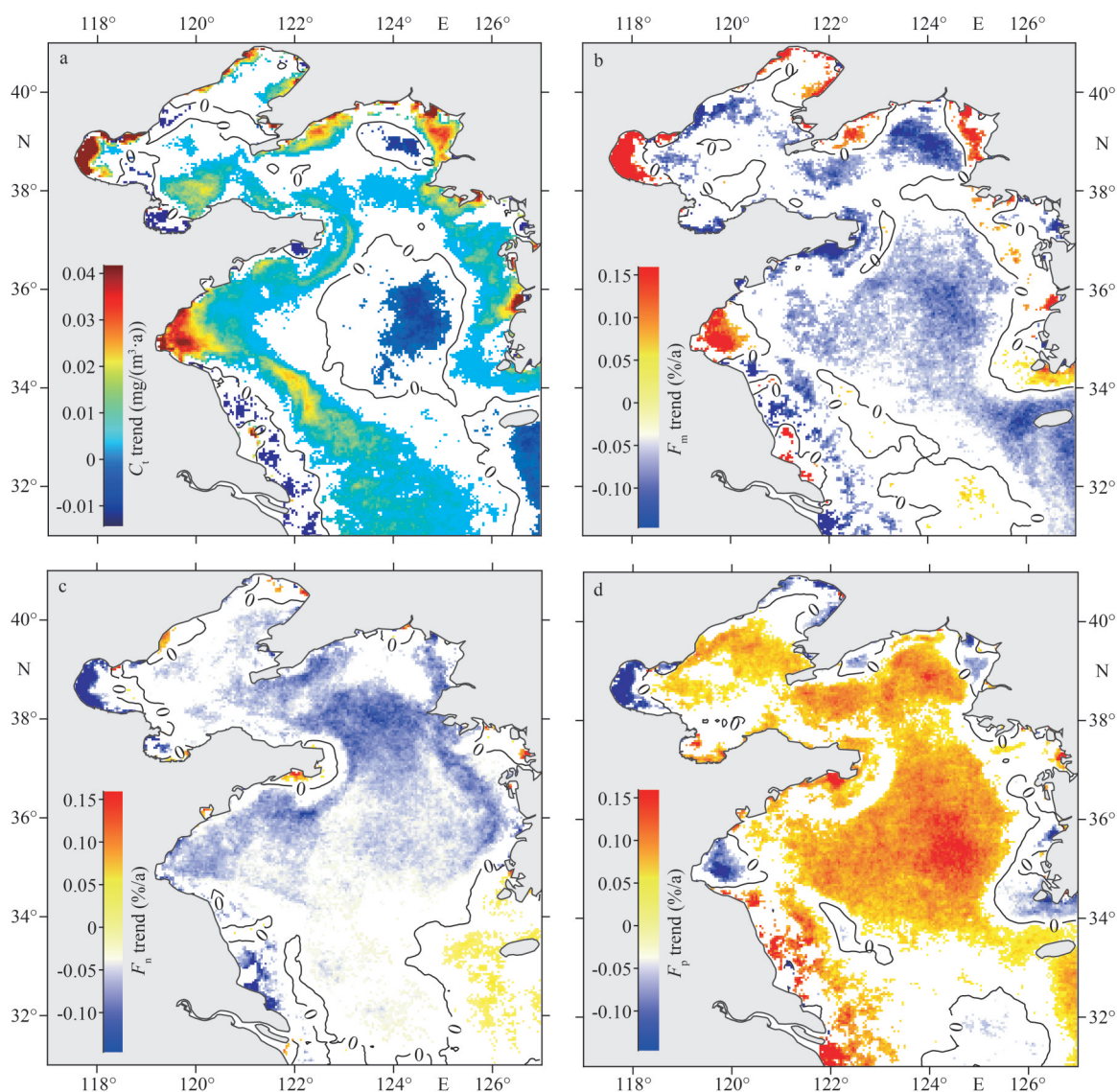


Fig.7 Spatial distributions of linear trends in PSC during 2003–2024

a. trends of total Chl- α concentration; b. trends of micro-phytoplankton contribution; c. trends of nano-phytoplankton contribution; d. trends of pico-phytoplankton contribution. Trend slopes were derived using Mann-Kendall (MK) trend analysis, with pixels showing statistically insignificant trends (at 95% confidence level) masked out.

45.33%, and 50.46% of the area presented significant trends in F_m , F_n , and F_p , respectively, in which negative trend dominated in F_m (31.78% of total area) and F_n (42.21%), whereas positive trends prevailed in F_p (45.80%). This contrasting spatial pattern between trends in large-sized and small-sized fractions suggests a potential functional shift in phytoplankton community structure within BYS, indicating a transition from larger to smaller phytoplankton functional groups.

Figure 8 presents the temporal variations and corresponding MK trends detected in the basin-averaged C_t and PSC time series for the entire BYS and its two sub-regions. Both the entire and coastal C_t showed statistically significant positive trends (0.005 and 0.009 mg/(m³·a), respectively), while offshore C_t displayed a non-significant negative trend. All F_m exhibited negative trends (-0.031%, -0.002 and -0.055%/a for the entire, coastal and offshore F_m , respectively), with only the trend of offshore F_m was statistically significant (Fig.8f). All

the F_n exhibited significant negative trends with the rates of -0.049%, -0.045% and -0.055%/a for the entire, coastal and offshore F_n , respectively (Fig.8g–i). Conversely, all the F_p increased (0.055%, 0.018% and 0.093%/a for the entire, coastal and offshore regions, respectively), though the coastal F_p trend was statistically non-significant (Fig.8k). These results indicate divergent phytoplankton biomass trends between coastal (increasing) and offshore (decreasing) waters, while revealing a consistent shift toward smaller phytoplankton size classes throughout the BYS.

3.5 Interpretation of PSC variations based on RF models

The SHAP method was applied to interpret the PSC-environment relationships, derived from RF models, with results visualized separately for the coastal (Fig.9) and offshore (Fig.10) province. All established RF models demonstrated satisfactory performance, explaining over 65% of the variation

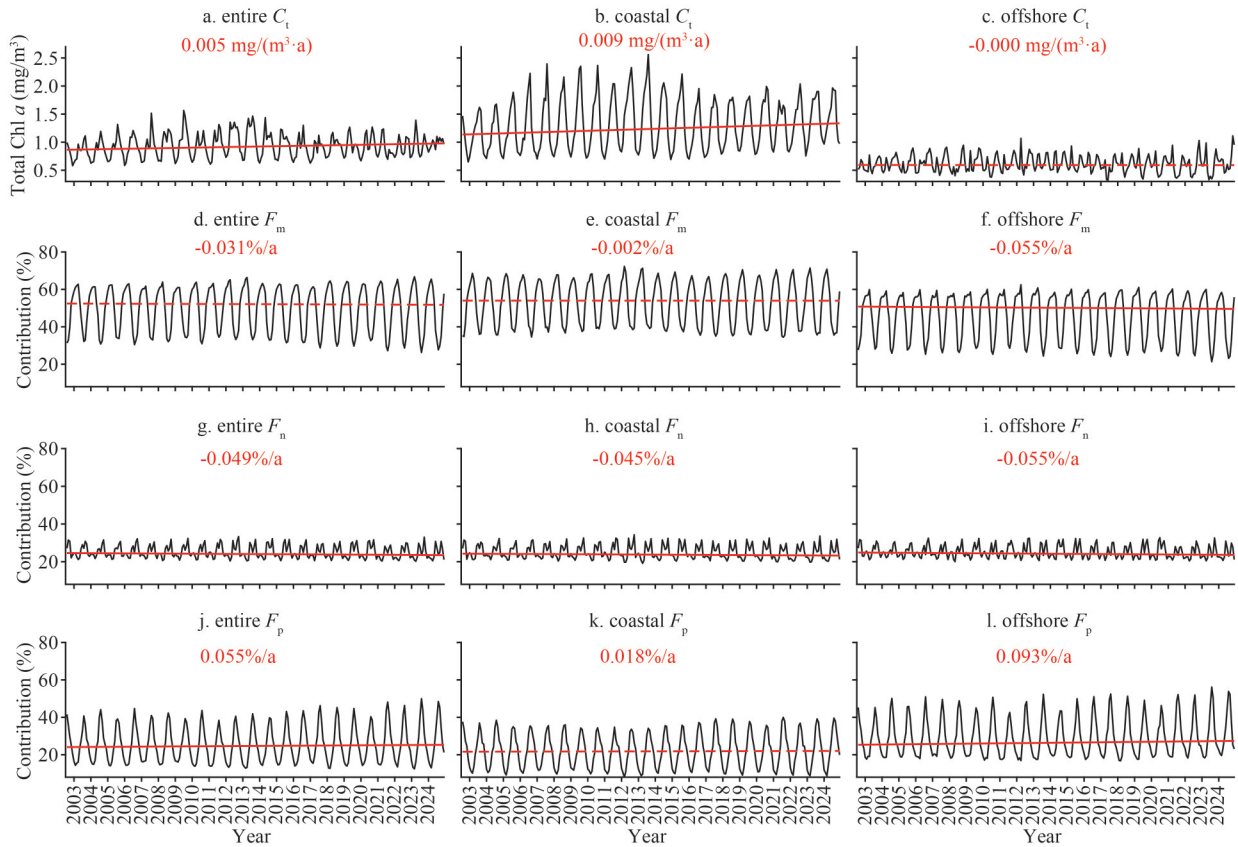


Fig.8 Temporal variations and MK trends of basin-averaged total Chl-*a* concentration and contribution of each size class in the entire BYS and its two sub-regions (as defined in Fig.1)

Gray solid lines represent original data series; red straight lines denote MK trends (solid for statistically significant trends, $P < 0.05$; dotted for non-significant trends, $P > 0.05$). C_t : total Chl-*a* concentration; F_m : micro-phytoplankton contribution; F_n : nano-phytoplankton contribution; F_p : pico-phytoplankton contribution.

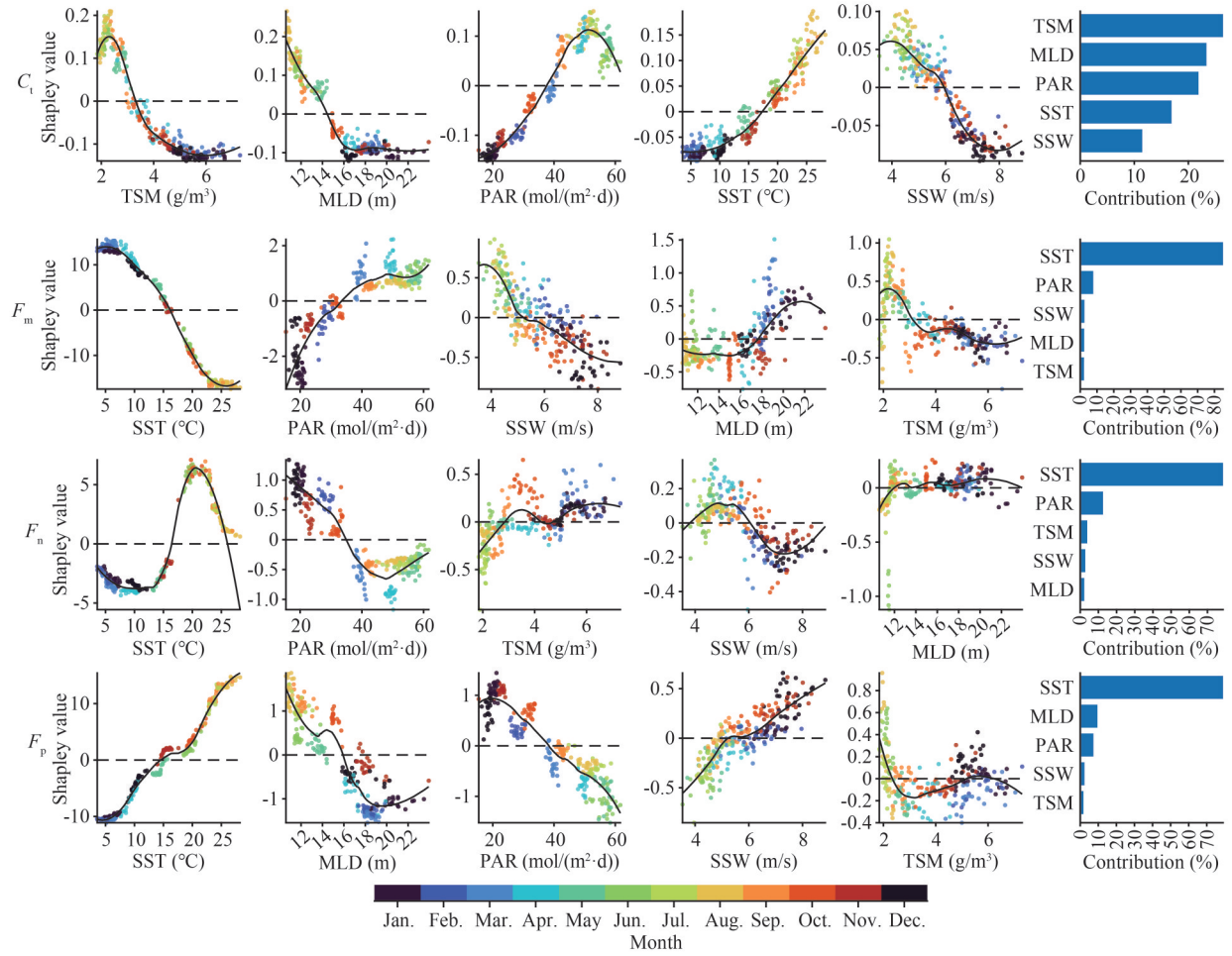


Fig.9 SHAP-based interpretation of PSC-environment relationships for coastal region using RF model

Left five columns show SHAP dependence plot and right column shows contributions of different variables ranked by aggregated absolute SHAP values. PSC: phytoplankton size class; RF: random forest; C_t : total Chl-*a* concentration; F_m : micro-phytoplankton contribution; F_n : nano-phytoplankton contribution; F_p : pico-phytoplankton contribution; SST: sea surface temperature; TSM: total suspended matter; PAR: photosynthetically active radiation; SSW: sea surface wind; MLD: mixed layer depth.

in both C_t and F_{size} variations. The SHAP summary plots illustrate the marginal effect of each environmental variable (scatter plots), with LOESS (Locally Estimated Scatterplot Smoothing)-smoothed trend lines (black curves) highlighting dominant patterns (a span parameter of 0.5 with quadratic polynomial fitting). Accompanying bar plots quantify the relative contribution of each driver using mean absolute SHAP values. In the dependence plots, sampling points are color-coded by month to identify seasonal patterns in these interactions.

3.5.1 Coastal water

In coastal waters, TSM (26%) was the most influential factor governing C_t variations, followed by MLD (23%), PAR (22%), SST (18%), and SSW (11%) (Fig.9). TSM, MLD, and SSW generally

exhibit negative correlations with C_t , and their impacts showed pronounced threshold behaviors: $\sim 3.5 \text{ g/m}^3$ for TSM, $\sim 15 \text{ m}$ for MLD, and $\sim 6 \text{ m/s}$ for SSW. For instance, TSM below 3.5 g/m^3 (predominantly in summer) were associated with increased C_t , whereas levels above this threshold (mainly in winter) correlated with suppressed biomass. The threshold behavior of TSM may reflect the indirect modulation of phytoplankton growth by suspended particulate matter under conditions of light and nutrient trade-off in turbid coastal waters. PAR and SST displayed significant positive correlations with C_t , with clear promotion threshold at $\sim 38 \text{ mol/(m}^2\cdot\text{d)}$ and $\sim 16^\circ\text{C}$, respectively. These thresholds align with the seasonal transition to conditions favorable for phytoplankton growth: higher light and temperature in late spring/summer.

For F_{size} , SST overwhelmingly the dominant driver across all size classes, with contributions exceeding 70%. A significant negative correlation existed between SST and F_m . SST below $\sim 16^\circ\text{C}$ strongly promoted F_m , whereas warmer temperature inhibited it. This pattern is consistent with the known preference of micro-phytoplankton for cooler, nutrient-replete conditions (Marañón et al., 2012), typically found in winter and early spring for the BYS. The relationship between SST and F_m was highly nonlinear: inhibitory below $\sim 16^\circ\text{C}$, promotional between ~ 16 and $\sim 21^\circ\text{C}$, with the effect weakening and eventually reversing to inhibition above $\sim 26^\circ\text{C}$. This complex response may reflect shifting competitive balances and grazing pressures across thermal regimes. SST positively correlated with F_p with a clear promotional threshold above $\sim 16^\circ\text{C}$, supporting the ecological principle that smaller phytoplankton exhibit a metabolic advantage in warmer, more

stratified waters (Marañón et al., 2012). PAR was the secondary driver for F_m and F_n , with higher PAR ($>40 \text{ mol}/(\text{m}^2\cdot\text{d})$) favoring F_m but suppressing F_n , potentially due to differing photoadaptive strategies. MLD influenced F_n , promoting it below $\sim 16 \text{ m}$ and inhibiting it in deeper mixed layers.

3.5.2 Offshore water

In offshore waters, TSM remained the primary driver of C_t variations (30% contribution), followed by PAR, SST, MLD, and SSW (contributions between 10%–20%) (Fig.10). All factors exhibited nonlinear impacts on C_t variations with identifiable threshold. TSM showed a tri-phasic effect on C_t : inhibitory below $\sim 0.8 \text{ g}/\text{m}^3$ (summer-dominant), promotional between ~ 0.8 and $\sim 1.8 \text{ g}/\text{m}^3$ (winter-dominant), and insignificant above $\sim 1.8 \text{ g}/\text{m}^3$. This suggests an optimal, moderate turbidity range that balances light availability with nutrient supply via resuspension. PAR promotes C_t below $50 \text{ mol}/(\text{m}^2\cdot\text{d})$

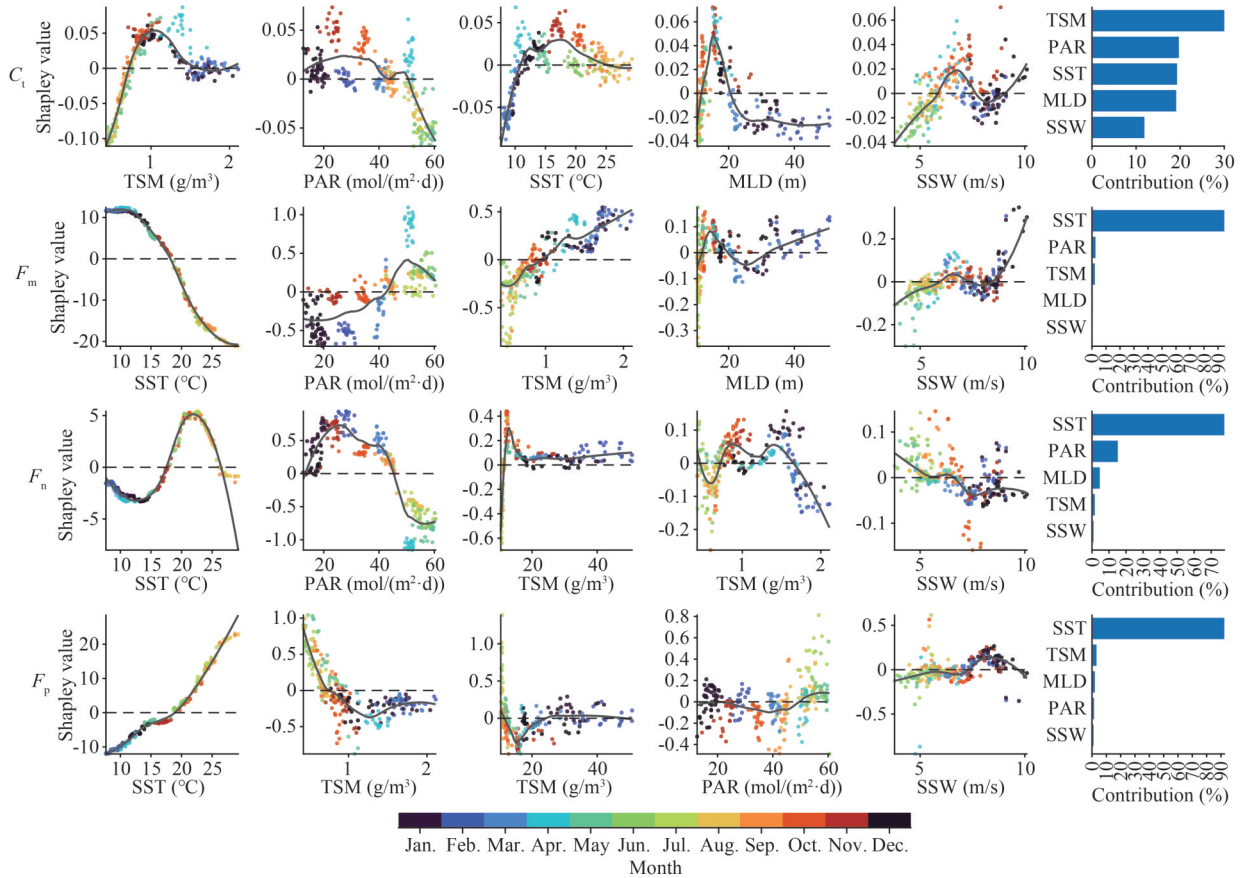


Fig.10 SHAP-based interpretation of PSC-environment relationships for offshore region using RF model

Left five columns show SHAP dependence plot and right column shows contributions of different variables ranked by aggregated absolute SHAP values. PSC: phytoplankton size class; RF: random forest; C_t : total Chl-*a* concentration; F_m : micro-phytoplankton contribution; F_n : nano-phytoplankton contribution; F_p : pico-phytoplankton contribution; SST: sea surface temperature; TSM: total suspended matter; PAR: photosynthetically active radiation; SSW: sea surface wind; MLD: mixed layer depth.

(autumn/winter) but inhibited it at higher intensities, possibly indicating photoinhibition or co-limitation by other factors in summer. SST promoted C_t within at optimal 10–25 °C but was inhibitory outside this range. MLD promoted C_t only within a narrow range of 10–20 m, highlighting the importance of moderate mixing for nutrient renewal without causing light limitation.

For F_{size} in offshore waters, SST dominated, explaining >70% of the variance in F_n , and >90% in F_m and F_p . The negative SST- F_m relationship showed a threshold at ~18 °C, above which micro-phytoplankton was strongly suppressed. The SST- F_n relationship mirrored the coastal pattern but with a higher promotional threshold (~18 °C), above which the effect weakened (~21 °C) and eventually reversed (~26 °C). The SST- F_p relationship was positively correlated with a threshold at ~18 °C. The exceptional dominance of SST underscores the fundamental role of temperature in structuring phytoplankton communities through its control on metabolic rates, stratification, and nutrient cycling—a manifestation of the “temperature-size rule” in pelagic ecosystems (Chen et al., 2012; Pálffy and Smeti, 2024). PAR was the secondary factor for F_m and F_n , with a threshold at ~42 mol/(m²·d), while TSM secondarily influenced F_n , promoting it below ~0.8 g/m³. SSW exhibited the weakest and most unstable influence on C_t and F_{size} in the offshore province.

4 DISCUSSION

This study integrated multi-year cruise-observations with satellite remote sensing to optimize an abundance-based PSC model for the BYS, to characterize spatiotemporal patterns in PSC, and to explore environmental drivers for PSC variability using an interpretable machine learning framework. The 22-year time series (2003–2024) generated through this study offers a critical baseline for detecting future climate change impacts on marine ecosystems. By bridging the gap between discrete shipboard measurements and continuous satellite observations, our approach enables a more comprehensive understanding of phytoplankton dynamics that is essential for ecosystem-based management in this rapidly changing region.

4.1 Validation of remote sensing model for PSC retrieval

The regional optimization and robust validation

of PSC models using extensive cruise data represents a significant advancement in monitoring marine ecosystem dynamics. Our study demonstrated that satellite-derived PSC estimates are in strong agreement with in-situ measurements. The spatiotemporal dynamics of satellite-derived PSC in the BYS, as delineated in this study, generally align with the PSC patterns observed in previous in-situ observations (Fu et al., 2009; Jiang et al., 2019) and other remotely sensed estimates (Sun et al., 2019b; Liu et al., 2021). Previous studies have reported significant variations in these parameters across global oceanic regions (Sun et al., 2023) and the marginal seas adjacent to mainland China (Sun et al., 2019b; Liu et al., 2021). Our study also identified distinct relationships between these parameters and SST, contrasting with the findings of previous studies in similar regions (Sun et al., 2019b). These discrepancies are likely due to different in-situ datasets used for parameter tuning in the models. In comparison, our study compiled an unprecedented volume of HPLC data specifically for the BYS, covering a wider temperatures range (approximately 0–30 °C) and a more dense spatiotemporal coverage than previous studies (Sun et al., 2019b; Liu et al., 2021). Consequently, the PSC model developed in this study may offer improved reliability and can be further applied for PSC dynamics analysis in the BYS.

4.2 Nearshore-offshore dichotomy in PSC climatology

The pronounced nearshore-offshore dichotomy in phytoplankton total biomass and size structure represents a fundamental characteristic of the BYS ecosystem. Our analysis revealed that coastal waters were consistently dominated by micro-phytoplankton, contributing 40%–60% of total biomass year-round. In contrast, offshore waters exhibited a clear seasonal alternation, with pico-phytoplankton dominating in summer and micro-phytoplankton prevailing during the rest of the year (Figs. 5 & 6). Under the influence of the East Asian monsoon climate and substantial anthropogenic inputs, coastal waters of the BYS remain generally well-mixed and eutrophic throughout the year (Wei et al., 2016). Consequently micro-phytoplankton persistently dominate these regions, with proportion exceeding 60% in the cold season (Fig. 6), reflecting the combined effects of high nutrient availability and mixing conditions that favor large-sized phytoplankton. In offshore waters, nutrient

conditions show marked seasonal variability (Jin et al., 2013), leading to a corresponding seasonal shift in PSC contributions (Fig.6). The water column in the central BYS is slightly stratified in the spring, strong stratified in summer and vertically mixed in both the autumn and winter (Wei et al., 2016). Associated with these hydrographic changes, surface nutrient become depleted in the surface layer during the summer and replenished in other seasons (Wang et al., 2003; Wei et al., 2010; Tan and Shi, 2012). As a result, micro-phytoplankton dominates the central BYS in autumn and winter, when nutrient supply to the upper layer is sufficient. Meanwhile, strong wintertime mixing, along with inputs from Yellow Sea Coastal Current, further accumulates nutrients in winter. From winter to spring, a pronounced spring phytoplankton bloom develops (Fig.6), peaking notably in April—a pattern widely reported in previous studies (Xuan et al., 2011; Zhou et al., 2013). Our size-fractionated results indicate that this bloom is supported by growth across all phytoplankton size classes, with micro-phytoplankton contributing the largest share. This spring dominance can be attributed to a combination of favorable early-spring conditions: high residual nutrient concentrations from winter mixing, gradually increasing solar radiation, the onset of seasonal stratification that retains cells in the euphotic zone, and relatively low zooplankton grazing pressure (Xuan et al., 2011; Jiang et al., 2019). Together, these factors provide a competitive advantage to micro-phytoplankton, allowing them to flourish in the offshore waters of BYS for several weeks in spring. As conditions shift from spring to summer, the dominant PSC in offshore waters transitions from micro- to pico-phytoplankton. This shift is likely driven by the competitive advantage of smaller phytoplankton under rising surface temperatures, enhancing stratification, and reducing nutrient levels (Jiang et al., 2019). Overall, the seasonal PSC patterns and their drivers identified here are consistent with findings from earlier studies in this region (Fu et al., 2009; Jiang et al., 2019).

4.3 Long-term phytoplankton trend: biomass increase and size miniaturization

Our analysis of the 22-years (2003–2024) satellite-derived PSC dataset revealed a significant long-term trend in the BYS phytoplankton: an overall increase in total biomass coupled with a distinct shift towards smaller size fractions (e.g., increasing dominance of picoplankton). The

decreased phytoplankton size of BYS is consistent with the well-known size-reduced phytoplankton changes in marginal china seas (Wang et al., 2023a) and global decline of phytoplankton size under global warming scenario (Lewandowska et al., 2014).

The RF modeling of C_t and F_{size} time series provided comprehensive insights into the environmental controls on the variations in phytoplankton total biomass and size structure. TSM emerged as the primary determinant of biomass variability, reflecting its role as a proxy for nutrient bioavailability. High TSM correlates with nutrient adsorption on particles, directly fueling primary productivity. In turbid waters, TSM also mediates light availability, indirectly favoring small taxa adapted to low-light conditions. The dual effect of TSM on phytoplankton biomass makes the correlations between TSM and C_t highly nonlinear where moderate TSM concentrations can stimulate phytoplankton growth while high concentrations become inhibitory (Figs.9 & 10). Critically, our Random Forest model pinpointed SST as the overwhelmingly dominant driver governing the proportional changes in size structure (i.e., the shift towards smaller fractions). This finding strongly supports the hypothesis that rising SST exerts primary control on phytoplankton size composition in this region. The mechanisms likely involve both direct physiological effects and indirect physical effects (Lee et al., 2019). The direct physiological effect was that smaller phytoplankton generally exhibit higher maximum growth rates and are better adapted to warmer conditions. The indirect physical effect was that warming intensifies vertical stratification, hindering the upward supply of nutrients from deeper layers. The resulting oligotrophic conditions in the surface layer favor smaller phytoplankton with superior nutrient affinity and competitive abilities under low-nutrient regimes, such as cyanobacteria (e.g., *Synechococcus*, *Prochlorococcus*) and small eukaryotes (Lee et al., 2023). The dominance of SST in controlling F_{size} underscores the importance of temperature-mediated physiological and community processes. Warmer temperatures favor picoplankton through several mechanisms (Pálffy and Smeti, 2024): (1) enhanced metabolic rates that advantage small cells with faster division cycles; (2) reformed community composition via temperature-dependent species sorting; (3) intensified stratification that reduces nutrient flux to surface waters; and (4) altered

grazing pressure as zooplankton communities respond to warming.

4.4 Uncertainty and future perspective

While our study provides comprehensive insights into the spatiotemporal dynamics of phytoplankton size structure in the BYS, several methodological limitations should be acknowledged. First, the in-situ PSC data was not directly measure phytoplankton cell size, but an approximation via HPLC-based pigment-to-size conversion. This conversion relies on the assumptions that specific taxonomic groups are associated with certain diagnostic pigments and that these groups consistently correspond to predefined size classes. However, some diagnostic pigments can be shared by phytoplankton groups spanning a wide size range and may vary with physiological state (Gholizadeh and Robeson, 2016; Leblanc et al., 2018), thus introducing uncertainty into the derived PSC estimates (Chase et al., 2020). Second, MODIS Chl-*a* retrieval in the optically complex and highly variable waters of the BYS, especially in turbid nearshore regions, remains challenging. Errors source from atmospheric correction, bottom reflectance in shallow areas, and deviations of bio-optical algorithms in Case-2 waters can introduce noise into the R_{rs} spectra, which can propagate uncertainties into the satellite PSC estimates. Third, HPLC data represents a surface “snapshot”, whereas MODIS captures a depth-integrated signal (over the first optical depth). This mismatch may lead to inconsistencies between in situ and satellite-derived PSC values.

The abundance-based PSC model used in this study is inherently an indirect empirical approach. Although it offers advantages in simplicity and operational feasibility, and achieves reliable retrieval accuracy, it is less grounded in bio-optical principles, less interpretable, and likely less transferable than more direct methods based on spectral features such as phytoplankton absorption coefficients. In the BYS region, previous studies have examined such direct absorption-based inversion methods with reliable results (Sun et al., 2019a). In fact, the accuracy of abundance-based model appears comparable to those reported for absorption-based approaches in BYS (Sun et al., 2019a, b; Liu et al., 2021). Nevertheless, a systematic quantitative comparison between these different methodologies was not conducted in the present work, and can be attempted in our future

researches.

Environmental dataset, including TSM, SST, SSW, MLD, and PAR were used to explore potential drivers of PSC variations. We acknowledge that in a complex coastal system such as the BYS, these satellite-derived or model simulated environmental products inevitably carry inherent uncertainties. However, since our analysis primarily relied on spatially averaged values over the study region, the influence of local-scale data errors might be partly mitigated. Furthermore, the interpretation focused on relative spatiotemporal patterns rather than absolute magnitudes, thereby reducing potential biases in results arising from inaccuracies in the input environmental data.

A methodological consideration arises from our use of a temperature-parameterized model to retrieve PSC, while independently identifying SST as the dominant ecological driver. This design could, in principle, introduce circularity if the conclusion were drawn directly from the retrieval model itself. However, the dominance of SST was determined through an independent, explainable machine-learning framework (Random Forest with SHAP analysis), where SST competed equally with other environmental predictors (TSM, PAR, MLD, SSW). The consistent, data-driven ranking of SST as the primary factor (>70%–90% contribution to size-fraction variance) supports its ecological primacy, irrespective of the retrieval model’s structure. Future studies employing alternative, non-temperature-based inversion methods (e.g., absorption-based approaches) in the region would provide a valuable independent test of this SST dependency.

The two-decade database presented here offers unprecedented spatiotemporal resolution for analyzing phytoplankton community dynamics in this important marginal-sea system. The methodological framework and ecological insights derived from this study hold relevance beyond the BYS. The integrated approach—combining long-term satellite observations with a regionally optimized PSC model and explainable machine-learning techniques for driver attribution—provides a transferable paradigm for assessing phytoplankton community shifts in other temperate coastal and shelf seas worldwide. The observed shift toward smaller phytoplankton under warming and enhanced stratification aligns with patterns reported in diverse ocean regions (Ding et al., 2023), suggesting that the processes identified here likely reflect a widespread ecological response to climate forcing (Brander and Kiørboe, 2020). The

phytoplankton minimization may reduce the trophic-transfer efficiency, weaken carbon export potential, and alter biogeochemical cycling (Finkel et al., 2010; Marañón, 2015). Such changes suggest a potential decline in the efficiency of biological carbon pump and a restructuring of pelagic food web-favoring micro-zooplankton grazers over direct consumption by larger copepods and fish larvae, which might ultimately affect fisheries productivity (Litchman et al., 2007). These findings establish a critical baseline for evaluating future climate change impacts on continental shelf ecosystems and highlight the need for adaptive management strategies that account for on-going phytoplankton community restructuring.

5 CONCLUSION

This study has constructed a novel, two-decade (2002–2024) remote sensing database of phytoplankton size structure in the BYS by optimizing an abundance-based PSC model using extensive in situ HPLC pigment data and MODIS-derived remote sensing reflectance. The result of this study establishes a critical baseline for understanding long-term phytoplankton community changes in the BYS. We demonstrate the power of integrating remote sensing, pigment analysis, optimized modeling, and explainable machine learning technique to unravel complex ecological drivers. The observed coastal-offshore gradient reaffirms the primacy of nutrient supply and light limitation in structuring phytoplankton communities. The basin-wide trends of increasing biomass (driven primarily by TSM-related processes) and decreasing cell size (driven primarily by warming and its effects on stratification and nutrient supply) might indicate fundamental ecological shifts. The movement towards a smaller-sized phytoplankton community has profound implications for the pelagic food web structure, potentially reducing the efficiency of energy transfer to higher trophic levels (e.g., fisheries) and altering biogeochemical cycles, particularly carbon export. Continued monitoring, refined modeling incorporating nutrient dynamics, and focused studies on trophic impacts are essential to fully understand and predict the future trajectory of this vital marine ecosystem under ongoing environmental change.

6 DATA AVAILABILITY STATEMENT

The satellite derived PSC data can be acquired

from <https://www.csdac.ac.cn/dataDetails/b71c0d02d1b2447b8e6cfe4d65911780>. The satellite-derived Chl-*a* data can be obtained from <https://www.csdac.ac.cn/dataDetails/501373119ba749158a9abcf1d0304273>. MODIS Aqua R_{rs} and PAR data were obtained from NASA Ocean Color (<https://oceancolor.gsfc.nasa.gov>). Daily Optimum Interpolation Sea Surface Temperature (OISSTv2) data were sourced from NOAA National Centers for Environmental Information (<https://www.ncei.noaa.gov/products/optimum-interpolation-sst>). MLD data were derived from the ARMOR3D global ocean reanalysis product, accessible via the Copernicus Marine Service (<https://doi.org/10.48670/moi-00019>). The other datasets will be made available on request.

References

- Agawin N S R, Duarte C M, Agustí S. 2000. Nutrient and temperature control of the contribution of picoplankton to phytoplankton biomass and production. *Limnology and Oceanography*, **45**(8): 1891–1891, <https://doi.org/10.4319/lo.2000.45.3.0591>.
- Aiken J, Pradhan Y, Barlow R et al. 2009. Phytoplankton pigments and functional types in the Atlantic Ocean: a decadal assessment, 1995–2005. *Deep Sea Research Part II: Topical Studies in Oceanography*, **56**(15): 899–917, <https://doi.org/10.1016/j.dsr2.2008.09.017>.
- Atlas R, Hoffman R N, Ardizzone J et al. 2011. A cross-calibrated, multiplatform ocean surface wind velocity product for meteorological and oceanographic applications. *Bulletin of the American Meteorological Society*, **92**(2): 157–174, <https://doi.org/10.1175/2010BAMS2946.1>.
- Barnett T P, Pierce D W, Schnur R. 2001. Detection of anthropogenic climate change in the world's oceans. *Science*, **292**(5515): 270–274, <https://doi.org/10.1126/science.1058304>.
- Brander K, Kiørboe T. 2020. Decreasing phytoplankton size adversely affects ocean food chains. *Global Change Biology*, **26**(10): 5356–5357, <https://doi.org/10.1111/gcb.15216>.
- Breiman L. 2001. Random forests. *Machine Learning*, **45**(1): 5–32, <https://doi.org/10.1023/A:1010933404324>.
- Brewin R J W, Ciavatta S, Sathyendranath S et al. 2017. Uncertainty in ocean-color estimates of chlorophyll for phytoplankton groups. *Frontiers in Marine Science*, **4**: 104, <https://doi.org/10.3389/fmars.2017.00104>.
- Brewin R J W, Hardman-Mountford N J, Lavender S J et al. 2011. An intercomparison of bio-optical techniques for detecting dominant phytoplankton size class from satellite remote sensing. *Remote Sensing of Environment*, **115**(2): 325–339, <https://doi.org/10.1016/j.rse.2010.09.004>.
- Brewin R J W, Sathyendranath S, Hirata T et al. 2010. A three-component model of phytoplankton size class for the Atlantic Ocean. *Ecological Modelling*, **221**(11): 1472–1483, <https://doi.org/10.1016/j.ecolmodel.2010.02.014>.

- Brewin R J W, Sathyendranath S, Jackson T et al. 2015. Influence of light in the mixed-layer on the parameters of a three-component model of phytoplankton size class. *Remote Sensing of Environment*, **168**: 437-450, <https://doi.org/10.1016/j.rse.2015.07.004>.
- Campbell J W. 1995. The lognormal distribution as a model for bio-optical variability in the sea. *Journal of Geophysical Research: Oceans*, **100**(C7): 13237-13254, <https://doi.org/10.1029/95JC00458>.
- Chase A P, Kramer S J, Haëntjens N et al. 2020. Evaluation of diagnostic pigments to estimate phytoplankton size classes. *Limnology and Oceanography: Methods*, **18**(10): 570-584, <https://doi.org/10.1002/lom3.10385>.
- Chen B Z, Landry M R, Huang B Q et al. 2012. Does warming enhance the effect of microzooplankton grazing on marine phytoplankton in the ocean? *Limnology and Oceanography*, **57**(2): 519-526, <https://doi.org/10.4319/lo.2012.57.2.0519>.
- de Boyer Montégut C, Madec G, Fischer A S et al. 2004. Mixed layer depth over the global ocean: an examination of profile data and a profile-based climatology. *Journal of Geophysical Research: Oceans*, **109**(C12): C12003, <https://doi.org/10.1029/2004JC002378>.
- de Lima Oliveira A, Rudorff N, Sathyendranath S et al. 2025. Phytoplankton size structure in a subtropical area from ocean colour and its applications. *Journal of Marine Systems*, **248**: 104036, <https://doi.org/10.1016/j.jmarsys.2024.104036>.
- Devred E, Sathyendranath S, Stuart V et al. 2011. A three component classification of phytoplankton absorption spectra: application to ocean-color data. *Remote Sensing of Environment*, **115**(9): 2255-2266, <https://doi.org/10.1016/j.rse.2011.04.025>.
- Ding X, Liu J X, Liu W W et al. 2023. Phytoplankton communities miniaturization driven by extreme weather in subtropical estuary under climate changes. *Water Research*, **245**: 120588, <https://doi.org/10.1016/j.watres.2023.120588>.
- Feurer M, Klein A, Eggensperger K et al. 2019. Auto-sklearn: efficient and robust automated machine learning. In: Hutter F, Kotthoff L, Vanschoren J eds. *Automated Machine Learning: Methods, Systems, Challenges*. Springer, Cham. p.113-134, https://doi.org/10.1007/978-3-030-05318-5_6.
- Field C B, Behrenfeld M J, Randerson J T et al. 1998. Primary production of the biosphere: integrating terrestrial and oceanic components. *Science*, **281**(5374): 237-240, <https://doi.org/10.1126/science.281.5374.237>.
- Finkel Z V, Beardall J, Flynn K J et al. 2010. Phytoplankton in a changing world: cell size and elemental stoichiometry. *Journal of Plankton Research*, **32**(1): 119-137, <https://doi.org/10.1093/plankt/fbp098>.
- Frouin R, Franz B A, Werdell P J. 2003. The SeaWiFS PAR product. In: Hooker S B, Firestone E R eds. *Algorithm Updates for the Fourth SeaWiFS Data Reprocessing*. p.46-50, <https://ntrs.nasa.gov/citations/20030102166>.
- Fu M Z, Wang Z L, Li Y et al. 2009. Phytoplankton biomass size structure and its regulation in the Southern Yellow Sea (China): seasonal variability. *Continental Shelf Research*, **29**(18): 2178-2194, <https://doi.org/10.1016/j.csr.2009.08.010>.
- Furuya K, Hayashi M, Yabushita Y et al. 2003. 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**(2): 367-387, [https://doi.org/10.1016/S0967-0645\(02\)00460-5](https://doi.org/10.1016/S0967-0645(02)00460-5).
- Gao Y J, Zhao J Y, Han L. 2023. Quantifying the nonlinear relationship between block morphology and the surrounding thermal environment using random forest method. *Sustainable Cities and Society*, **91**: 104443, <https://doi.org/10.1016/j.scs.2023.104443>.
- Gholizadeh H, Robeson S M. 2016. Revisiting empirical ocean-colour algorithms for remote estimation of chlorophyll-*a* content on a global scale. *International Journal of Remote Sensing*, **37**(11): 2682-2705, <https://doi.org/10.1080/01431161.2016.1183834>.
- Gilbert R O. 1987. *Statistical Methods for Environmental Pollution Monitoring*. Van Nostrand Reinhold Co., New York, <https://doi.org/10.2307/2531935>.
- Gittings J A, Brewin R J W, Raitos D E et al. 2019. Remotely sensing phytoplankton size structure in the Red Sea. *Remote Sensing of Environment*, **234**: 111387, <https://doi.org/10.1016/j.rse.2019.111387>.
- Guidi L, Stemann L, Jackson G A et al. 2009. Effects of phytoplankton community on production, size, and export of large aggregates: a world-ocean analysis. *Limnology and Oceanography*, **54**(6): 1951-1963, <https://doi.org/10.4319/lo.2009.54.6.1951>.
- Hao Q, Chai F, Xiu P et al. 2019. Spatial and temporal variation in chlorophyll *a* concentration in the Eastern China Seas based on a locally modified satellite dataset. *Estuarine, Coastal and Shelf Science*, **220**: 220-231, <https://doi.org/10.1016/j.ecss.2019.01.004>.
- Hirata T, Hardman-Mountford N J, Brewin R J W et al. 2011. Synoptic relationships between surface Chlorophyll-*a* and diagnostic pigments specific to phytoplankton functional types. *Biogeosciences*, **8**(2): 311-327, <https://doi.org/10.5194/bg-8-311-2011>.
- Hirsch R M, Slack J R. 1984. A nonparametric trend test for seasonal data with serial dependence. *Water Resources Research*, **20**(6): 727-732, <https://doi.org/10.1029/WR020i006p00727>.
- Hirsch R M, Slack J R, Smith R A. 1982. Techniques of trend analysis for monthly water quality data. *Water Resources Research*, **18**(1): 107-121, <https://doi.org/10.3133/wfr81488>.
- Hoegh-Guldberg O, Bruno J F. 2010. The impact of climate change on the world's marine ecosystems. *Science*, **328**(5985): 1523-1528, <https://doi.org/10.1126/science.1189930>.
- IOCCG. 2014. *Phytoplankton functional types from Space*. IOCCG, Dartmouth, <https://epic.awi.de/id/eprint/36000/>.
- Jiang Z B, Chen J F, Gao Y X et al. 2019. Regulation of spatial changes in phytoplankton community by water column stability and nutrients in the Southern Yellow

- Sea. *Journal of Geophysical Research: Biogeosciences*, **124**(8): 2610-2627, <https://doi.org/10.1029/2018JG004785>.
- Jin J, Liu S M, Ren J L et al. 2013. Nutrient dynamics and coupling with phytoplankton species composition during the spring blooms in the Yellow Sea. *Deep Sea Research Part II: Topical Studies in Oceanography*, **97**: 16-32, <https://doi.org/10.1016/j.dsr2.2013.05.002>.
- Leblanc K, Quéguiner B, Diaz F et al. 2018. Nanoplanktonic diatoms are globally overlooked but play a role in spring blooms and carbon export. *Nature Communications*, **9**(1): 953, <https://doi.org/10.1038/s41467-018-03376-9>.
- Lee D G, Oh J H, Noh K M et al. 2023. What controls the future phytoplankton change over the Yellow and East China Seas under global warming? *Frontiers in Marine Science*, **10**: 1010341, <https://doi.org/10.3389/fmars.2023.1010341>.
- Lee K H, Jeong H J, Lee K et al. 2019. Effects of warming and eutrophication on coastal phytoplankton production. *Harmful Algae*, **81**: 106-118, <https://doi.org/10.1016/j.hal.2018.11.017>.
- Lewandowska A M, Boyce D G, Hofmann M et al. 2014. Effects of sea surface warming on marine plankton. *Ecology Letters*, **17**(5): 614-623, <https://doi.org/10.1111/ele.12265>.
- Lin L, Liu D Y, Wang Y Q et al. 2023. Effect of wind on summer chlorophyll-*a* variability in the Yellow Sea. *Frontiers in Marine Science*, **9**: 1104258, <https://doi.org/10.3389/fmars.2022.1104258>.
- Litchman E, Klausmeier C A, Schofield O M et al. 2007. The role of functional traits and trade-offs in structuring phytoplankton communities: scaling from cellular to ecosystem level. *Ecology Letters*, **10**(12): 1170-1181, <https://doi.org/10.1111/j.1461-0248.2007.01117.x>.
- Liu D Y, Wang Y Q. 2013. Trends of satellite derived chlorophyll-*a* (1997-2011) in the Bohai and Yellow Seas, China: effects of bathymetry on seasonal and inter-annual patterns. *Progress in Oceanography*, **116**: 154-166, <https://doi.org/10.1016/j.pocean.2013.07.003>.
- Liu H R, Liu X, Xiao W P et al. 2021. Spatial and temporal variations of satellite-derived phytoplankton size classes using a three-component model bridged with temperature in Marginal Seas of the Western Pacific Ocean. *Progress in Oceanography*, **191**: 102511, <https://doi.org/10.1016/j.pocean.2021.102511>.
- Lundberg S M, Lee S I. 2017. A unified approach to interpreting model predictions. In: Proceedings of the 31st International Conference on Neural Information Processing Systems. Curran Associates Inc., Long Beach, USA. p.4768-4777, <https://doi.org/10.48550/arXiv.1705.07874>.
- Lv T, Liu D Y, Zhou P et al. 2022. The coastal front modulates the timing and magnitude of spring phytoplankton bloom in the Yellow Sea. *Water Research*, **220**: 118669, <https://doi.org/10.1016/j.watres.2022.118669>.
- Marañón E. 2015. Cell size as a key determinant of phytoplankton metabolism and community structure. *Annual Review of Marine Science*, **7**: 241-264, <https://doi.org/10.1146/annurev-marine-010814-015955>.
- Marañón E, Cermeño P, Latasa M et al. 2012. Temperature, resources, and phytoplankton size structure in the ocean. *Limnology and Oceanography*, **57**(5): 1266-1278, <https://doi.org/10.4319/lo.2012.57.5.1266>.
- Marañón E, Cermeño P, Latasa M et al. 2015. Resource supply alone explains the variability of marine phytoplankton size structure. *Limnology and Oceanography*, **60**(5): 1848-1854, <https://doi.org/10.1002/lno.10138>.
- Moore T S, Brown C W. 2020. Incorporating environmental data in abundance-based algorithms for deriving phytoplankton size classes in the Atlantic Ocean. *Remote Sensing of Environment*, **240**: 111689, <https://doi.org/10.1016/j.rse.2020.111689>.
- Pálffy K, Smeti E. 2024. Combined effect of warming, nutrients, and species pool size on the seasonal variability of phytoplankton composition: a modeling perspective. *Limnology and Oceanography*, **69**(5): 1056-1069, <https://doi.org/10.5194/os-12-379-2016>.
- Pitarch J, Volpe G, Colella S et al. 2016. Remote sensing of chlorophyll in the Baltic Sea at basin scale from 1997 to 2012 using merged multi-sensor data. *Ocean Science*, **12**(2): 379-389, <https://doi.org/10.5194/os-12-379-2016>.
- Reynolds R W, Smith T M. 1994. Improved global sea surface temperature analyses using optimum interpolation. *Journal of Climate*, **7**(6): 929-948, [https://doi.org/10.1175/1520-0442\(1994\)007<0929:IGSSTA>2.0.CO;2](https://doi.org/10.1175/1520-0442(1994)007<0929:IGSSTA>2.0.CO;2).
- Reynolds R W, Smith T M, Liu C et al. 2007. Daily high-resolution-blended analyses for sea surface temperature. *Journal of Climate*, **20**(22): 5473-5496, <https://doi.org/10.1175/2007JCLI1824.1>.
- Sahay A, Ali S M, Gupta A et al. 2017. Ocean color satellite determinations of phytoplankton size class in the Arabian Sea during the winter monsoon. *Remote Sensing of Environment*, **198**: 286-296, <https://doi.org/10.1016/j.rse.2017.06.017>.
- Saulquin B, Fablet R, Mangin A et al. 2013. Detection of linear trends in multisensor time series in the presence of autocorrelated noise: application to the chlorophyll-*a* SeaWiFS and MERIS data sets and extrapolation to the incoming Sentinel 3-OLCI mission. *Journal of Geophysical Research: Oceans*, **118**(8): 3752-3763, <https://doi.org/10.1002/jgrc.20264>.
- Seegers B N, Stumpf R P, Schaeffer B A et al. 2018. Performance metrics for the assessment of satellite data products: an ocean color case study. *Optics Express*, **26**(6): 7404-7422, <https://doi.org/10.1364/OE.26.007404>.
- Sen P K. 1968. Estimates of the regression coefficient based on kendall's tau. *Journal of the American Statistical Association*, **63**(324): 1379-1389, <https://doi.org/10.1080/01621459.1968.10480934>.
- Sieburth J M, Smetacek V, Lenz J. 1978. Pelagic ecosystem structure: heterotrophic compartments of the plankton and their relationship to plankton size fractions. *Limnology and Oceanography*, **23**(6): 1256-1263, <https://doi.org/10.1002/lno.10138>.

- doi.org/10.4319/lo.1978.23.6.1256.
- Smetacek V, Cloern J E. 2008. On phytoplankton trends. *Science*, **319**(5868): 1346-1348, <https://doi.org/10.1126/science.1151330>.
- Sun D Y, Huan Y, Qiu Z F et al. 2017. Remote-sensing estimation of phytoplankton size classes from GOCI satellite measurements in Bohai Sea and Yellow Sea. *Journal of Geophysical Research: Oceans*, **122**(10): 8309-8325, <https://doi.org/10.1002/2017JC013099>.
- Sun D Y, Huan Y, Wang S Q et al. 2019a. Remote sensing of spatial and temporal patterns of phytoplankton assemblages in the Bohai Sea, Yellow Sea, and east China sea. *Water Research*, **157**: 119-133, <https://doi.org/10.1016/j.watres.2019.03.081>.
- Sun X R, Brewin R J W, Sathyendranath S et al. 2023. Coupling ecological concepts with an ocean-colour model: phytoplankton size structure. *Remote Sensing of Environment*, **285**: 113415, <https://doi.org/10.1016/j.rse.2022.113415>.
- Sun X R, Shen F, Brewin R J W et al. 2019b. Twenty-year variations in satellite-derived chlorophyll-*a* and phytoplankton size in the Bohai Sea and Yellow Sea. *Journal of Geophysical Research: Oceans*, **124**(12): 8887-8912, <https://doi.org/10.1029/2019JC015552>.
- Sun X R, Shen F, Liu D Y et al. 2018. In situ and satellite observations of phytoplankton size classes in the entire continental shelf sea, China. *Journal of Geophysical Research: Oceans*, **123**(5): 3523-3544, <https://doi.org/10.1029/2017JC013651>.
- Tan S C, Shi G Y. 2012. The relationship between satellite-derived primary production and vertical mixing and atmospheric inputs in the Yellow Sea cold water mass. *Continental Shelf Research*, **48**: 138-145, <https://doi.org/10.1016/j.csr.2012.07.015>.
- Trees C C, Clark D K, Bidigare R R et al. 2000. Accessory pigments versus chlorophyll *a* concentrations within the euphotic zone: a ubiquitous relationship. *Limnology and Oceanography*, **45**(5): 1130-1143, <https://doi.org/10.4319/lo.2000.45.5.1130>.
- Turner K J, Mouw C B, Hyde K et al. 2021. Optimization and assessment of phytoplankton size class algorithms for ocean color data on the Northeast U.S. continental shelf. *Remote Sensing of Environment*, **267**: 112729, <https://doi.org/10.1016/j.rse.2021.112729>.
- Uitz J, Huot Y, Bruyant F et al. 2008. Relating phytoplankton photophysiological properties to community structure on large scales. *Limnology and Oceanography*, **53**(2): 614-630, <https://doi.org/10.4319/lo.2008.53.2.0614>.
- Vantrepotte V, Mélin F. 2011. Inter-annual variations in the SeaWiFS global chlorophyll *a* concentration (1997-2007). *Deep Sea Research Part I: Oceanographic Research Papers*, **58**(4): 429-441, <https://doi.org/10.1016/j.dsr.2011.02.003>.
- Victoria A H, Maragatham G. 2021. Automatic tuning of hyperparameters using Bayesian optimization. *Evolving Systems*, **12**(1): 217-223, <https://doi.org/10.1007/s12530-020-09345-2>.
- Wang B D, Wang X L, Zhan R. 2003. Nutrient conditions in the Yellow Sea and the East China Sea. *Estuarine, Coastal and Shelf Science*, **58**(1): 127-136, [https://doi.org/10.1016/S0272-7714\(03\)00067-2](https://doi.org/10.1016/S0272-7714(03)00067-2).
- Wang F, Li X G, Tang X H et al. 2023a. The seas around China in a warming climate. *Nature Reviews Earth & Environment*, **4**(8): 535-551, <https://doi.org/10.1038/s43017-023-00453-6>.
- Wang Y Q, Gao Z Q. 2020. Contrasting chlorophyll-*a* seasonal patterns between nearshore and offshore waters in the Bohai and Yellow Seas, China: a new analysis using improved satellite data. *Continental Shelf Research*, **203**: 104173, <https://doi.org/10.1016/j.csr.2020.104173>.
- Wang Y Q, Liu D Y, Gao Z Q et al. 2023b. Characterizing spatial patterns of satellite-derived chlorophyll-*a* in the Bohai and Yellow Seas of China using self-organizing maps (SOM) approach. *Marine Pollution Bulletin*, **193**: 115176, <https://doi.org/10.1016/j.marpolbul.2023.115176>.
- Wang Y Q, Liu D Y, Tang D L. 2017. Application of a generalized additive model (GAM) for estimating chlorophyll-*a* concentration from MODIS data in the Bohai and Yellow Seas China. *International Journal of Remote Sensing*, **38**(3): 639-661, <https://doi.org/10.1080/01431161.2016.1268733>.
- Wang Y Q, Liu D Y, Wang Y J et al. 2019. Evaluation of standard and regional satellite chlorophyll-*a* algorithms for moderate-resolution imaging spectroradiometer (MODIS) in the Bohai and Yellow Seas, China: a comparison of chlorophyll-*a* magnitude and seasonality. *International Journal of Remote Sensing*, **40**(13): 4980-4995, <https://doi.org/10.1080/01431161.2019.1577579>.
- Wang Y Q, Liu D Y, Wang Y J et al. 2024. Optimizing an abundance-based model for satellite remote sensing of phytoplankton size classes in the Bohai and Yellow Seas of China. *IEEE Transactions on Geoscience and Remote Sensing*, **62**: 4204813, <https://doi.org/10.1109/TGRS.2024.3383391>.
- Wang Y Q, Tian X P, Gao Z Q. 2021. Evolution of satellite derived chlorophyll-*a* trends in the Bohai and Yellow Seas during 2002-2018: comparison between linear and nonlinear trends. *Estuarine, Coastal and Shelf Science*, **259**: 107449, <https://doi.org/10.1016/j.ecss.2021.107449>.
- Wei Q S, Fu M Z, Ge R F et al. 2010. Chemicohydrographic characteristics and the seasonal variations of nutrients at 35°N transect in the cold water mass of the Southern Yellow Sea. *Environmental Science*, **31**(9): 2603-2074, <https://doi.org/10.1631/jzus.A1000244>.
- Wei Q S, Yu Z G, Wang B D et al. 2016. Coupling of the spatial-temporal distributions of nutrients and physical conditions in the southern Yellow Sea. *Journal of Marine Systems*, **156**: 30-45, <https://doi.org/10.1016/j.jmarsys.2015.12.001>.
- Xu Y, Zhang D, Lin J et al. 2024. Prediction of phytoplankton biomass and identification of key influencing factors using interpretable machine learning models. *Ecological Indicators*, **158**: 111320, <https://doi.org/10.1016/j.ecolind.2023.111320>.
- Xuan J L, Zhou F, Huang D et al. 2011. Physical processes

- and their role on the spatial and temporal variability of the spring phytoplankton bloom in the central Yellow Sea. *Acta Ecologica Sinica*, **31**(1): 61-70, <https://doi.org/10.1016/j.chnaes.2010.11.011>.
- Yamaguchi H, Kim H C, Son Y B et al. 2012. Seasonal and summer interannual variations of SeaWiFS chlorophyll *a* in the Yellow Sea and East China Sea. *Progress in Oceanography*, **105**: 22-29, <https://doi.org/10.1016/j.pocean.2012.04.004>.
- Zapata M, Rodríguez F, Garrido J L. 2000. Separation of chlorophylls and carotenoids from marine phytoplankton: a new HPLC method using a reversed phase C₈ column and pyridine-containing mobile phases. *Marine Ecology Progress Series*, **195**: 29-45, <https://doi.org/10.3354/meps195029>.
- Zhai F G, Liu Z Z, Gu Y Z et al. 2023. Satellite-observed interannual variations in sea surface chlorophyll-*a* concentration in the Yellow Sea over the past two decades. *Journal of Geophysical Research: Oceans*, **128**(4): e2022JC019528, <https://doi.org/10.1029/2022JC019528>.
- Zhai F G, Wu W F, Gu Y Z et al. 2021. Interannual-decadal variation in satellite-derived surface chlorophyll-*a* concentration in the Bohai Sea over the past 16 years. *Journal of Marine Systems*, **215**: 103496, <https://doi.org/10.1016/j.jmarsys.2020.103496>.
- Zhang M W, Tang J W, Dong Q et al. 2010. Retrieval of total suspended matter concentration in the Yellow and East China Seas from MODIS imagery. *Remote Sensing of Environment*, **114**(2): 392-403, <https://doi.org/10.1016/j.rse.2009.09.016>.
- Zhou F, Xuan J L, Huang D J et al. 2013. The timing and the magnitude of spring phytoplankton blooms and their relationship with physical forcing in the central Yellow Sea in 2009. *Deep Sea Research Part II: Topical Studies in Oceanography*, **97**: 4-15, <https://doi.org/10.1016/j.dsr2.2013.05.001>.