Journal of Oceanology and Limnology

https://doi.org/10.1007/s00343-026-5464-2

Research Paper

Regionalizing the vertically generalized production model (VGPM) for Sentinel-2 estimation of phytoplankton primary productivity in turbid inland waters*

Ting SONG[1], Yanjuan XU[1], Hujun ZHANG[1], Bingchuan ZHU[1], Ling LI[1], Dongyang CHENG[3], Ge LIU[2, **]

[1] *Jiangsu Wuxi Environmental Monitoring Center, Wuxi 214121, China*

[2] *State Key Laboratory of Black Soils Conservation and Utilization, Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Changchun 130102, China*

[3] *Suzhou Chenyang Hengrui Information Technology Co., Ltd., Suzhou 215000, China*

Received Dec. 2, 2025; accepted in principle Feb. 11, 2026; accepted for publication Mar. 9, 2026

Abstract Phytoplankton primary productivity (PP) is a key indicator of carbon fixation and ecosystem functioning in inland waters. Using high-resolution Sentinel-2 imagery, this study developed a regionalized Vertically Generalized Production Model (VGPM) optimized for highly turbid, shallow lakes to estimate modeled PP dynamics in Taihu Lake from 2019 to 2025. The multi-year mean daily PP in Taihu Lake ranged from 820.68 to 1 527.11 mg C/(m²·d), with a lake-wide average of ~1 232.28 mg C/(m²·d). Spatially, modeled PP decreased from the bays to the central basin and from west to east, with Zhushan and Meiliang bays maintaining the highest productivity (up to ~1 667.93 mg C/(m²·d)). At the same time, the eastern region dominated by aquatic vegetation showed lower values (<950 mg C/(m²·d)). Temporally, modeled PP exhibited a strong seasonal cycle, with higher values in summer and autumn and lower values in winter and spring, forming a bimodal pattern with peaks in May and September (>1 800 mg C/(m²·d)). Moderate water temperature and sufficient light availability promoted productivity, whereas turbidity (light attenuation) limited phytoplankton growth. By integrating corrections for phytoplankton vertical distribution, algal bloom classification, and the removal of aquatic vegetation signals, the model's performance was greatly improved. The Google Earth Engine (GEE)-based workflow enabled efficient long-term modeled PP monitoring and provides a transferable framework for quantifying carbon cycling, guiding eutrophic lake restoration, and assessing climate responses.

Keyword: Sentinel-2; Taihu Lake; primary productivity (PP); Vertically Generalized Production Model (VGPM) model

1 INTRODUCTION

Phytoplankton primary productivity (PP), defined as the rate of carbon (C) fixation by phytoplankton, underpins the global carbon cycle in aquatic ecosystems. As the principal primary producers in inland and marine waters, phytoplankton form the foundation of aquatic food webs and drive the biological productivity of water bodies. Inland waters, through the interplay of air-water gas exchange, terrestrial carbon input, and sediment burial, function simultaneously as carbon sources, sinks, and reservoirs, thereby playing a crucial role in sustaining the stability of aquatic ecosystems.

Lakes, despite covering only a small fraction of the ocean's surface area, serve as vital hubs for carbon transport, transformation, and long-term storage. They are actively involved in a suite of multiscale biogeochemical processes, ranging from cross-boundary fluxes of dissolved inorganic and organic carbon to CO_2 and CH_4 exchange at the

* Supported by the National Key Research and Development Program of China (No. 2021YFB3901101), the National Natural Science Foundation of China (No. 42171385), the Open Fund of the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University (No. 22R02), and the Youth Innovation Promotion Association of the Chinese Academy of Sciences (No. 2022228)

** Corresponding author: liuge@iga.ac.cn

water-atmosphere interface and carbon burial at the sediment-water boundary (Tranvik et al., 2009). These heat and gas exchanges not only modulate watershed microclimates, but also alter regional radiation and energy balances, introducing a long-overlooked source of uncertainty in global carbon accounting (Cole et al., 2007).

Recent estimates indicate that total phytoplankton productivity in global lakes has reached ~1.77 Pg C/a, underscoring the substantial role of phytoplankton in mediating carbon fixation across land-water interfaces (Gao et al., 2021). As a fundamental metric of ecosystem functioning, PP is shaped by both biological factors (e.g., community composition, biomass) and environmental conditions (e.g., nutrient availability, light, temperature, and turbidity) (Anderson et al., 2021), resulting in pronounced spatial and temporal heterogeneity (Zhang et al., 2017). Importantly, PP constitutes a primary source of new organic carbon and energy input within lake ecosystems, fueling ecological processes and trophic dynamics (Kauer et al., 2015).

Accurately capturing the spatiotemporal variability of PP is therefore essential for elucidating material and energy flows in freshwater systems. Such monitoring is critical not only for quantifying lake carbon budgets but also for assessing ecosystem responses to climate change and informing effective water-quality and carbon management strategies.

Phytoplankton PP can be measured directly through in situ observations or estimated indirectly using semi-empirical models. Common in situ techniques include isotope tracer methods based on ^{13}C or ^{14}C incorporation (Cassar et al., 2011), the dissolved oxygen method (Halsey et al., 2013), and fluorescence-based approaches (Silsbe et al., 2015). While these methods can provide relatively accurate PP estimates, they are labor-intensive and time-consuming, limiting their applicability for capturing large-scale and long-term spatiotemporal dynamics.

In contrast, satellite remote sensing offers a practical and cost-effective alternative to traditional field-based monitoring, with clear advantages in spatial coverage, temporal resolution, and revisit frequency (Ye et al., 2015). Over past decades, researchers have developed various empirical, semi-empirical, and analytical models to estimate primary production from remote sensing data (Platt, 1986; Behrenfeld and Falkowski, 1997; Pérez et al., 2005). Among these, the VGPM proposed by Behrenfeld and Falkowski (1997) has been widely applied due to its relatively low input requirements. Notably, two key input variables—chlorophyll-*a* concentration and euphotic depth—can be derived directly from satellite observations, enabling large-scale estimation of aquatic primary productivity via remote sensing.

Numerous studies have successfully integrated remote sensing data with VGPM to estimate oceanic phytoplankton productivity. For instance, Shih et al. (2021) evaluated the performance and limitations of multiple PP retrieval models, including VGPM, by combining in situ and satellite datasets. Westberry et al. (2023) reviewed the evolution of global PP modeling in the ocean-color satellite era, highlighting the central role of VGPM in advancing the understanding of ocean productivity. Gregg et al. (2003) revealed decadal trends in global marine PP and their associations with climate variability.

However, the VGPM framework was initially developed for open ocean (Case 1) environments, where water optical properties are primarily influenced by phytoplankton and the inherent optical assumptions generally hold. Applying this framework directly to optically complex inland waters is challenging. In Case 2 waters, constituents such as suspended sediments, colored dissolved organic matter (CDOM), and detritus disrupt the relationship between satellite reflectance and key bio-optical parameters, including chlorophyll-*a* and euphotic depth. This disruption introduces significant uncertainty in parameter retrieval and, consequently, PP estimates. In optically complex inland waters—particularly those with high turbidity—accurately retrieving photosynthetic parameters from satellite imagery remains challenging due to limited optical signal penetration and poor water clarity (Palmer et al., 2015).

Taihu Lake, with a surface area of approximately 2 338 km^2, a maximum depth of 3 m, and an average depth of 1.89 m, has a relative shallowness index (defined as the square root of lake area divided by mean depth) of 25.6 (Bachmann et al., 2000), classifying it as a typical turbid, Type-II shallow lake. As both a critical regional carbon sink and a hotspot for algal blooms, phytoplankton PP in Taihu Lake plays a central role in regulating carbon cycling, driving phytoplankton succession, and sustaining energy flow within the aquatic ecosystem. However, due to its optically complex waters and frequent meteorological disturbances, conventional VGPM frameworks—originally developed for oceanic environments—often yield substantial uncertainties when applied to such highly turbid inland waters. Developing regionally adaptive remote-sensing methods for PP estimation is therefore essential to improve the accuracy and reliability of carbon cycle assessments in inland aquatic systems.

In this study, we integrate a comprehensive literature review with in situ observations to establish remote-sensing inversion models for chlorophyll-*a* concentration and euphotic depth, tailored to the unique environmental characteristics of Taihu Lake. By incorporating bloom dynamics and ERA5 meteorological reanalysis data, we further account for phytoplankton vertical distribution patterns, enabling refined estimation of integrated water-column biomass. Based on this, a localized VGPM model is developed to better accommodate the lake's high turbidity conditions, thereby enabling more accurate PP retrieval. A key innovation of this work is the systematic analysis of the spatiotemporal patterns and driving mechanisms of phytoplankton PP in turbid shallow lakes.

The objectives of this study are fourfold: (1) to develop robust remote-sensing algorithms for chlorophyll *a* and euphotic depth specific to Taihu Lake, enabling accurate retrieval of key bio-optical parameters; (2) to optimize integrated phytoplankton biomass estimation by incorporating vertical distribution characteristics and meteorological drivers, thereby refining the structure of the regionalized VGPM model; (3) to build a fast, reproducible, and operational PP monitoring workflow for Taihu Lake on the Google Earth Engine (GEE) platform, enabling high spatiotemporal resolution tracking; and (4) to characterize the spatiotemporal dynamics of phytoplankton primary productivity in Taihu Lake from 2019 to 2025, and to elucidate its responses to climatic variability and nutrient fluctuations.

2 MATERIAL AND METHOD

2.1 Study area

Taihu Lake, the third-largest freshwater lake in China, is situated between 119°54′E and 120°36′E and 30°56′N and 31°31′N in the Changjiang (Yangtze) River Delta. It represents a vital lacustrine ecosystem in this region. As a typical shallow lake, Taihu Lake has long been plagued by eutrophication and frequent cyanobacterial blooms, resulting in pronounced ecological challenges. Due to spatial heterogeneity in hydrodynamic conditions, water quality, and anthropogenic pressures, the lake exhibits marked ecological zonation (Liu et al., 2025a). The eastern region is dominated by extensive macrophyte coverage (characteristic of a grass-type, macrophyte-dominated lake), while the remaining areas are primarily phytoplankton-dominated (algae-type ecosystem with frequent blooms) (Liu et al., 2025b).

This coexistence of macrophyte-dominated and algae-dominated zones presents a significant challenge for remote sensing-based PP estimation. In particular, the high spectral similarity between macrophytes and phytoplankton can lead to systematic overestimation of chlorophyll-*a* concentrations during satellite retrievals, ultimately biasing PP assessments (Hou et al., 2024). To address this issue, the present study employs an ecological zoning approach, delineating Taihu Lake into macrophyte-dominated and bloom-dominated regions (Fig.1). In macrophyte-rich zones, aquatic vegetation signals are explicitly identified and excluded from the analysis, enabling the more accurate extraction of phytoplankton-specific optical signatures and enhancing the relevance of PP estimates. In addition, based on the lake's geography and environmental conditions, this study divides Taihu Lake into seven subregions: East Taihu Lake, Zhushan Bay, Central Lake Area, Coastal Zone, Meiliang Bay, Gonghu Bay, and Xukou Bay (Fig.1).

2.2 Data acquisition

2.2.1 Satellite data

This study utilized atmospherically corrected

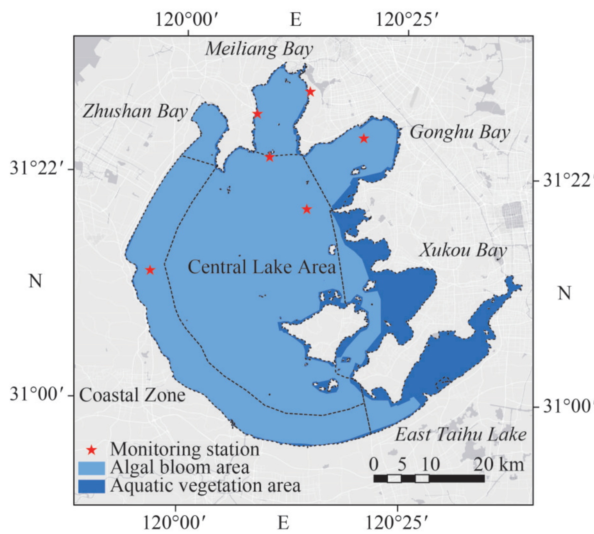

Fig.1 Location of Taihu Lake and its ecological subregions

The lake is delineated into macrophyte-dominated (grass-type) and phytoplankton bloom-dominated (algae-type) zones. Seven subregions are indicated: East Taihu Lake, Zhushan Bay, Central Lake Area, Coastal Zone, Meiliang Bay, Gonghu Bay, and Xukou Bay.

Sentinel-2 Multispectral Instrument (MSI) Level-2A products accessed via the GEE platform. These datasets offer 10-m spatial resolution and a five-day revisit interval, featuring 13 spectral bands that span the visible, near-infrared (VNIR), and shortwave infrared (SWIR) regions. Such characteristics make Sentinel-2 imagery well-suited for capturing fine-scale spatial and temporal dynamics of inland water systems. Table 1 summarizes the annual revisit frequency and the number of Sentinel-2 scenes acquired over the Taihu Lake region from 2019 through 2025.

2.2.2 Meteorological data

We utilized the ERA5-Land reanalysis dataset to provide meteorological inputs for the VGPM model and to investigate the drivers of PP variability. ERA5-Land, derived from the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA5 reanalysis (land component), offers high spatial and temporal resolution, enabling a precise depiction of fine-scale climatic variability relevant to inland aquatic ecosystems. The meteorological variables obtained include lake_mix_layer_temperature (LMLT), surface_solar_radiation_down wards_sum ($SSDR_{sum}$), total_precipitation (Precip), and 10-m wind components (U for east-west and V for north-south).

2.2.3 Surface water quality data

Water quality data were obtained from six automatic floating monitoring stations distributed across Taihu Lake, operated by the Jiangsu Provincial Environmental Monitoring Center (Fig.1). The monitored parameters included water surface temperature (WST), turbidity (in NTU), electrical conductivity (EC), dissolved oxygen (DO), ammonia nitrogen (NH_3-N), total nitrogen (TN), and total phosphorus (TP). Monthly averaged data from all six stations over the period 2019–2025 were compimay have contributed to identify the primary environmental drivers influencing PP.

Table 1 Annual revisit frequency and number of Sentinel-2 images over Taihu Lake (2019–2025)

Year	Revisit day	Number of images
2019	73	294
2020	73	298
2021	73	316
2022	73	314
2023	73	292
2024	72	290
2025	85	342

2.3 Method

2.3.1 Methodology and analysis framework

All satellite preprocessing, environmental data integration, and spatial analyses were performed within the GEE cloud-computing platform to ensure transparency and reproducibility (Fig.2). Sentinel-2 Level-2A surface reflectance imagery was acquired from the *COPERNICUS/S2_SR_HARMONIZED* collection and converted to surface reflectance using the standard scaling factor.

Cloud and cirrus contamination were reduced using a multi-stage masking strategy tailored for optically complex inland waters. Pixels classified as clouds, cirrus, or cloud shadows in the Sentinel-2 Scene Classification Layer (SCL) were first removed. Residual cloud contamination was then identified using spectral thresholds applied to the blue (B2) and red (B4) bands. The combined mask was expanded through morphological dilation to minimize cloud-edge effects and thin cirrus artifacts, and all subsequent analyses were restricted to cloud-free pixels.

For each target date, all Sentinel-2 scenes intersecting Taihu Lake within a one-day temporal window were collected and composited into a daily image using the mean of available cloud-free observations. Pixels without valid observations remained masked, and no gap filling was applied. Based on the ecological zonation of Taihu Lake, the study area was divided into macrophyte-dominated and phytoplankton-dominated (bloom-prone) regions. Chlorophyll-*a* retrieval and productivity estimation were conducted independently within each ecological zone and subsequently merged to produce a lake-wide dataset. Environmental variables, including wind speed, temperature, surface radiation, and photoperiod, were integrated at daily resolution within GEE. Detailed parameterization of the productivity model is described in Sections 2.3.2–2.3.4.

2.3.2 Parameterized model for PP estimation

Behrenfeld and Falkowski (1997) combined mechanistic principles of phytoplankton photosynthesis with empirical relationships to develop the VGPM. The model was developed using 1 698 observations collected between 1971 and 1994, spanning latitudes from 80°N to 70°S and

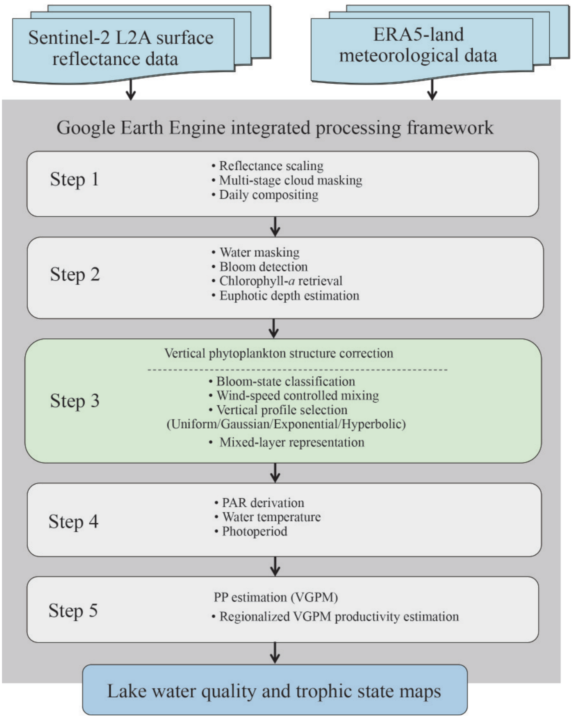

Fig.2 Workflow implemented in Google Earth Engine (GEE) for estimating daily phytoplankton primary productivity (PP) in Taihu Lake

The framework integrates satellite preprocessing, multi-stage cloud masking, bio-optical retrievals, environmental driver integration, vertical phytoplankton structure correction, and regionalized VGPM productivity estimation within a unified GEE environment to generate lake-wide daily PP maps.

encompassing both Case I and Case II waters. After normalizing chlorophyll concentration, photoperiod, and euphotic depth, they demonstrated that the vertical distribution of primary productivity follows a consistent functional form across diverse aquatic environments, establishing the theoretical foundation of the VGPM.

In a subsequent validation study, Behrenfeld and Falkowski compared model outputs with in situ measurements and showed that the VGPM accounted for 79% of the observed spatiotemporal variability in primary productivity (n=10 857) (Behrenfeld and Falkowski, 1997). In its simplified formulation, daily depth-integrated primary productivity within the euphotic zone is expressed as a function of phytoplankton biomass, euphotic depth, and surface irradiance (Eq.1).

$$\mathrm{PP}_{\mathrm{eu}} = 0.66125 \times P^{\mathrm{B}}_{\mathrm{opt}} \times \frac{E_0}{E_0 + 4.1} \times Z_{\mathrm{eu}} \times C_{\mathrm{opt}} \times D_{\mathrm{irr}}, \tag{1}$$

where $\mathrm{PP}_{\mathrm{eu}}$ is the daily depth-integrated primary productivity from the surface to the euphotic depth (mg C/(m^2·d)); $P^{\mathrm{B}}_{\mathrm{opt}}$ is the maximum chlorophyll-specific carbon fixation rate as a function of water temperature T; Z_{eu} denotes the euphotic depth (m, defined as the depth at which irradiance is reduced to 1% of its surface value); C_{opt} is the chlorophyll-*a* concentration (μg/L) in the water column; D_{irr} is the photoperiod (daylight duration, h); and E_0 is the daily cumulative surface photosynthetically active radiation (PAR, mol photons/(m^2·d)). The coefficients 0.661 25 and 4.1 are standard VGPM constants (Behrenfeld and Falkowski, 1997) and were not refitted in this study.

Among the VGPM input variables, daily surface photosynthetically active radiation (PAR; E_0, mol photons/(m^2·d)) was derived from ERA5-Land $\mathrm{SSDR}_{\mathrm{sum}}$ (J/(m^2·d)) using a two-step conversion.

First, total shortwave radiation was converted to the photosynthetically active fraction (400–700 nm) using a constant coefficient f_{PAR}=0.45, consistent with standard bio-optical modeling practice. Second, radiative energy (J) was converted to photon flux (mol photons) using an energy-to-quanta conversion factor α=4.62×10^{-6} mol photons/J (equivalent to 0.217 MJ/mol photons).

The conversion is expressed as:

$$E_0 = f_{\mathrm{PAR}} \times \mathrm{SSDR}_{\mathrm{sum}} \times \alpha, \tag{2}$$

where $\mathrm{SSDR}_{\mathrm{sum}}$ is daily total shortwave radiation (J/(m^2·d)), f_{PAR} is dimensionless, and α has units of mol photons/J. The resulting E_0 is therefore expressed in mol photons/(m^2·d).

Treating f_{PAR} as a constant introduces uncertainty, because the PAR fraction can vary with atmospheric conditions (e.g., cloud cover and aerosol loading) and solar zenith angle. Reported values typically range from approximately 0.45 to 0.50 under natural conditions (Tsubo and Walker, 2005). Variability in f_{PAR} propagates linearly into E_0 and nonlinearly into PP through the irradiance term $\frac{E_0}{E_0+4.1}$. The impact of this uncertainty is quantified in the sensitivity analysis (Section 4.1.2).

In our implementation, Z_{eu} is obtained from the empirical model described in Section 2.3.2; C_{opt} (chlorophyll-*a* concentration) is retrieved using the method in Section 2.3.3; D_{irr} is calculated from the geographic coordinates and date of observation; and $P^{\mathrm{B}}_{\mathrm{opt}}$ is estimated via a temperature-dependent function (Eq.3). In Eq.3, the water temperature T (used to derive $P^{\mathrm{B}}_{\mathrm{opt}}$ is obtained from the ERA5-Land LMLT product.

$$P_{\mathrm{opt}}^{B}\begin{cases}1.13 & (T\leq -1.0)\\ 4.00 & (T\geq 28.5)\\ P_{\mathrm{opt}}^{B} & (-1.0<T<28.5)\end{cases}, \quad (3)$$

where $P_{\mathrm{opt}}^{\mathrm{B}}$=1.2956+2.749×10$^{-1}T$+6.17×10$^{-2}T^2$–2.05×10$^{-2}T^3$+2.462×10$^{-3}T^4$–1.348×10$^{-4}T^5$+3.4132×10$^{-6}T^6$–3.27×10$^{-8}T^7$.

To evaluate the suitability of ERA5-Land LMLT as the temperature input for the productivity model, we compared LMLT with in situ daily mean water surface temperature (TW/WST) measured at a representative buoy station during 2019–2025. ERA5-Land LMLT was spatially collocated with the buoy location and temporally matched at a daily resolution. Model performance was assessed using ordinary least squares regression (slope, intercept, and R^2) and error metrics, including root mean square error (RMSE) and mean absolute percentage error (MAPE).

ERA5-Land LMLT closely reproduces the observed daily variability in water temperature (Fig.3). The comparison yielded 2 534, RMSE=0.65 K, MAPE=0.18%, and R^2=0.995, indicating excellent agreement between modeled and in situ measurements.

To quantify the propagation of temperature uncertainty into PP, we evaluated the sensitivity of the temperature-dependent term $P_{\mathrm{opt}}^{B}(T)$ to a ±0.65 K perturbation across a representative lake temperature range (285–300 K). The resulting relative change in Modeled PP was typically 0.8%–1.5%, with a maximum effect of <2% under cooler conditions.

These results demonstrate that the sub-degree temperature error introduces only minor uncertainty into PP estimates. Consequently, temperature-related uncertainty is unlikely to affect the dominant seasonal or interannual spatiotemporal patterns of productivity derived from the VGPM framework.

2.3.3 Empirical model for estimating Z_{eu}

According to the Lambert-Beer law, PAR irradiance attenuates exponentially with increasing water depth. The euphotic depth, Z_{eu}, represents the maximum depth to which sufficient light penetrates to support photosynthesis. It can be expressed as:

$$Z_{\mathrm{eu}}(\mathrm{PAR})=\frac{2\ln 10}{K_{\mathrm{d}}(\mathrm{PAR})}=\frac{4.605}{K_{\mathrm{d}}(\mathrm{PAR})}, \quad (4)$$

where Z_{eu} is the diffuse attenuation coefficient for PAR (/m). Based on in situ observations from Taihu Lake, Zhang et al. (2021) found a significant inverse correlation between water transparency and light attenuation (K_{d}) (n=81, R^2=0.98, P<0.001). The empirical relationship is given by:

$$K_{\mathrm{d}}(\mathrm{PAR})=0.896\times K_{\mathrm{d}}(490)^{0.873}. \quad (5)$$

Building on in situ spectral measurements from Taihu Lake, Shen et al. (2017) constructed a red-green band model for estimating K_{d}(490) from Sentinel-2 imagery. By incorporating the spectral response functions of the Sentinel-2 MSI sensor and converting to equivalent surface reflectance, they derived a robust model with high predictive accuracy (n=72, R^2=0.74, RMSE=0.86 /m, MAPE=21.05%):

$$K_{\mathrm{d}}(\mathrm{PAR})=13.23\times\frac{R_{\mathrm{rs}}(\mathrm{red})}{R_{\mathrm{rs}}(\mathrm{green})}-4.51. \quad (6)$$

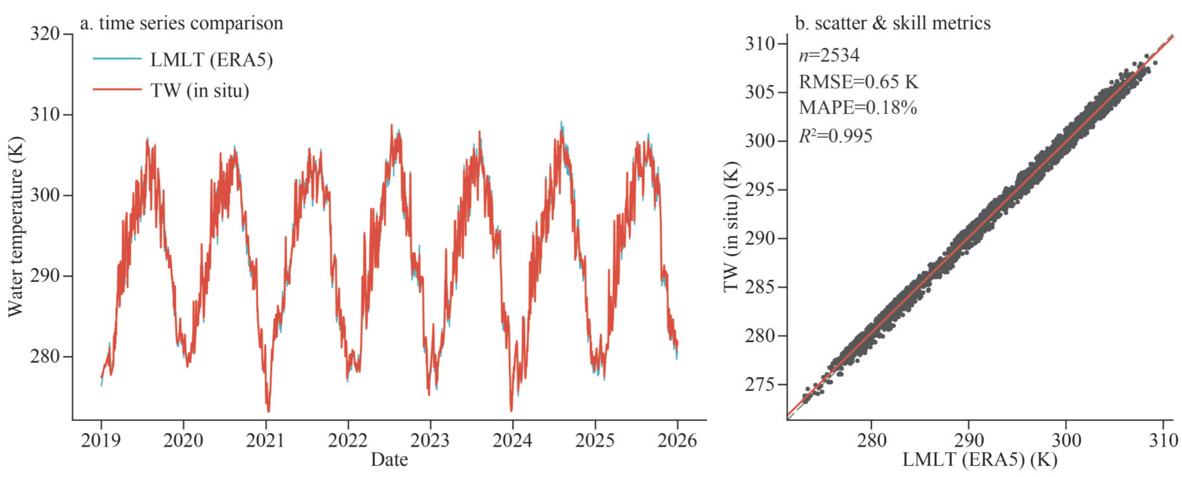

Fig.3 Evaluation of ERA5-Land LMLT against in situ daily mean water surface temperature (TW/WST) for 2019–2025

a. time series comparison demonstrating consistent seasonal and interannual variability between ERA5-Land LMLT and observations; b. catterplot of matched daily records (n=2 534) with corresponding performance metrics: RMSE=0.65 K, MAPE=0.18%, and R^2=0.995. The red solid line represents the linear regression fit, highlighting the strong agreement between ERA5-Land LMLT and observed water temperature.

In this model, R_{rs}(B4) and R_{rs}(B3) correspond to the remote sensing reflectance of Sentinel-2 bands B4 (665 nm, red) and B3 (560 nm, green), respectively. Combining Eqs.4–6, the estimation of Z_{eu} follows a chained dependency relationship: R_{rs}→ K_d(490)→K_d(PAR)→Z_{eu}. Through this chain, satellite-derived reflectance is first translated into spectral attenuation, then into broadband PAR attenuation, and finally into the euphotic depth. Accordingly, the euphotic depth Z_{eu}(PAR) can be estimated directly from Sentinel-2 reflectance data as:

$$Z_{eu}(\mathrm{PAR}) = \frac{5.14}{(13.23 \times \dfrac{R_{rs}(\mathrm{red})}{R_{rs}(\mathrm{green})} - 4.51)^{0.873}}. \quad (7)$$

Taihu Lake is a shallow, optically complex inland lake in which visible-band reflectance ratios can be affected by bottom reflectance and land adjacency effects. In optically shallow regions, lakebed reflection and shoreline mixing may contribute to the water-leaving radiance, thereby introducing bias into attenuation retrievals. To mitigate these effects, Z_{eu} estimation was restricted to optically deep open-water pixels using the water-masking procedures described in Section 2.3.4. Pixels influenced by macrophytes, emergent vegetation, or shoreline adjacency were excluded through a combined application of NDVI and SWIR-assisted water indices, reducing contamination from benthic reflection and adjacent land prior to applying the K_d-Z_{eu} relationship.

The chained empirical formulation introduces uncertainty propagation from reflectance retrieval to light-field estimation. Because Z_{eu} is inversely proportional to K_d(PAR), uncertainties in attenuation coefficients directly affect estimates of light penetration depth. To quantify this effect, a deterministic sensitivity analysis was performed by perturbing K_d(490) within the reported model uncertainty range (RMSE=0.86 /m). The perturbed values were propagated through Eqs.5 & 4 to recompute K_d(PAR) and Z_{eu}, and relative deviations from baseline estimates were evaluated. Under typical optical conditions in Taihu Lake, uncertainty in K_d(490) resulted in approximately 15%–25% relative variability in Z_{eu}. Despite this variability, the spatial distribution of underwater light availability remained stable and did not alter the dominant spatial patterns used for productivity estimation. Because primary productivity in the VGPM framework is jointly controlled by chlorophyll-*a* concentration, irradiance, temperature, and vertical structure correction, uncertainty associated with Z_{eu} is unlikely to dominate overall PP uncertainty. These results indicate that the derived Z_{eu} provides a physically consistent and robust optical input for the VGPM productivity model.

2.3.4 Estimation of chlorophyll-*a* concentration and vertical distribution

2.3.4.1 Remote sensing inversion of chlorophyll-*a* concentration

Following the spatial zoning strategy (Fig.1), Lake Taihu was divided into bloom-prone open-water regions and macrophyte-dominated regions. Distinct chlorophyll-*a* inversion models were developed for each zone. Before chlorophyll-*a* retrieval, we implemented a multi-index screening framework to distinguish water, land, and vegetation and to reduce misclassification in highly turbid waters and shoreline mixed pixels. We combined NDVI with SWIR-assisted water indices because NDVI alone can approach zero or become slightly positive in sediment-laden or mixed pixels, causing true water pixels to be erroneously excluded when NDVI is used as the sole discriminator. SWIR-based indices (e.g., MNDWI) and composite metrics (e.g., AWEI) are less sensitive to vegetation and built-up contamination, thereby improving masking robustness in optically complex and mixed-pixel environments.

In the phytoplankton-dominated areas, cyanobacterial blooms pixels were first identified using a combination of the Normalized Difference Vegetation Index (NDVI) and the Improved Floating Algae Index (IFAI). Pixels were classified as “bloom” when NDVI>0.1 and IFAI>0.03. These two indices are defined as follows:

$$\mathrm{NDVI} = \frac{R_{rs}(\mathrm{nir}) - R_{rs}(\mathrm{red})}{R_{rs}(\mathrm{nir}) + R_{rs}(\mathrm{red})}, \quad (8)$$

$$\mathrm{IFAI} = R_{rs}(\mathrm{green}) - R_{rs}(\mathrm{red}) + (R_{rs}(\mathrm{rede1}) - R_{rs}(\mathrm{green})) \times \frac{665-560}{704-560}, \quad (9)$$

where R_{rs}(nir) and R_{rs}(rede1) refer to the remote-sensing reflectance of Sentinel-2 bands B8 (833 nm) and B5 (705 nm), respectively.

To validate the bloom detection thresholds, an independent accuracy assessment was conducted using visually interpreted reference samples. Classification performance for cyanobacterial blooms and aquatic vegetation was evaluated using a confusion matrix. A total of 759 validation samples were analyzed, comprising 478 true positives, 254 true negatives, 26 false positives, and 19 false negatives. The resulting precision, recall, F1-score, and intersection-over-union were 0.948, 0.962, 0.955, and 0.914, respectively. These results demonstrate that the NDVI-IFAI threshold combination provides robust and reliable discrimination between bloom and non-bloom conditions, supporting its application for region-specific chlorophyll-*a* retrieval.

After bloom identification, the study area was separated into bloom and non-bloom zones. In situ chlorophyll-*a* measurements (2019–2025) from the Jiangsu Wuxi Environmental Monitoring Center were matched with concurrent Sentinel-2 observations to build empirical inversion models for both zones.

For the bloom region, a chlorophyll-*a* retrieval model was constructed using a red-edge band ratio (BR), defined as:

$$\mathrm{BR}=\frac{R_{\mathrm{rs}}(\mathrm{rede1})}{R_{\mathrm{rs}}(\mathrm{red})}. \tag{10}$$

An empirical regression of BR against in situ chlorophyll *a* yielded the following model for bloom conditions:

$$\mathrm{Chl}\ a_1=32.415\times\mathrm{BR}^2+3.5119\times\mathrm{BR}-10.47. \tag{11}$$

Model performance was assessed using an independent validation dataset. The bloom validation set comprised n=31 samples, yielding a coefficient of determination R^2 of 0.652, a RMSE of 34.02 μg/L, and a MAPE of 16.66% (Fig.4). As shown in Fig.4, the model successfully reproduced observed chlorophyll-*a* variability under high-biomass bloom conditions, demonstrating good agreement between satellite-derived estimates and in situ measurements.

In the non-bloom region, the chlorophyll-*a* model was derived from the Floating Band Absorption index (FBA), which incorporates spectral features from the blue, green, and coastal aerosol bands. The index is defined as:

$$\mathrm{FBA}=\left(\frac{1}{R_{\mathrm{rs}}(\mathrm{blue})}-\frac{0.2}{R_{\mathrm{rs}}(\mathrm{aerosol})}-\frac{0.8}{R_{\mathrm{rs}}(\mathrm{green})}\right)\times R_{\mathrm{rs}}(\mathrm{rede1}), \tag{12}$$

where R_{rs}(blue) and R_{rs}(aerosol) correspond to Sentinel-2 bands B2 (492 nm, blue) and B1 (443 nm, coastal aerosol), respectively. The FBA values were fitted to in situ chlorophyll-*a* observations, resulting in the empirical model for non-bloom conditions:

$$\mathrm{Chl}\ a_2=213.35\times\mathrm{FBA}+5.692. \tag{13}$$

Independent validation using n=86 samples yielded a coefficient of determination (R^2) of 0.613, a RMSE of 14.09 μg/L, and a MAPE of 37.89%. As shown in Fig.4, the model effectively captured chlorophyll-*a* variability across low-to-moderate concentration ranges characteristic of non-bloom conditions.

In the macrophyte-dominated zone (East Taihu Lake), aquatic vegetation can strongly alter visible-NIR reflectance and bias chlorophyll-*a* retrieval. Therefore, we applied a cross-validated open-water screening that integrates NDVI with SWIR-assisted water indices rather than relying on NDVI alone. We first excluded pixels with evident emergent or submerged vegetation signals (NDVI≥0). For the remaining pixels, we retained only those meeting a water-likelihood criterion based on the Normalized Difference Water Index (NDWI), the Modified NDWI (MNDWI), and, where needed, the Automated Water Extraction Index without shadow ($\mathrm{AWER}_{\mathrm{nsh}}$) to further suppress shoreline adjacency effects and mixed-pixel noise:

$$\mathrm{NDWI}=\frac{R_{\mathrm{rs}}(\mathrm{green})-R_{\mathrm{rs}}(\mathrm{nir})}{R_{\mathrm{rs}}(\mathrm{green})+R_{\mathrm{rs}}(\mathrm{nir})}, \tag{14}$$

$$\mathrm{MNDWI}=\frac{R_{\mathrm{rs}}(\mathrm{green})-R_{\mathrm{rs}}(\mathrm{swir1})}{R_{\mathrm{rs}}(\mathrm{green})+R_{\mathrm{rs}}(\mathrm{swir1})}, \tag{15}$$

$$\begin{aligned}\mathrm{AWEI}_{\mathrm{nsh}}=4\times(R_{\mathrm{rs}}(\mathrm{green})-R_{\mathrm{rs}}(\mathrm{swir1}))-\\(0.25\times R_{\mathrm{rs}}(\mathrm{nir})+2.75\times R_{\mathrm{rs}}(\mathrm{swir2})),\end{aligned} \tag{16}$$

where R_{rs}(swir1) and R_{rs}(swir2) correspond to Sentinel-2 bands B11 (SWIR1, 1 610 nm) and B12 (SWIR2, 2 190 nm), respectively. This multi-index screening reduces the likelihood that highly turbid water or shoreline mixed pixels are erroneously excluded by a single NDVI threshold and improves the stability of chlorophyll-*a* retrieval in optically complex nearshore waters. Pixels passing the open-water screening were treated as non-bloom water, and chlorophyll *a* was estimated using the non-bloom model (Eq.13).

2.3.4.2 Vertical distribution correction of phytoplankton

A combination of meteorological, hydrological, and biological processes regulates the vertical distribution of phytoplankton. In shallow lakes, wind-induced disturbances play a particularly prominent role in shaping phytoplankton distribution patterns (Ndong et al., 2014). The VGPM framework typically assumes a vertically uniform chlorophyll-*a* profile and uses surface concentration

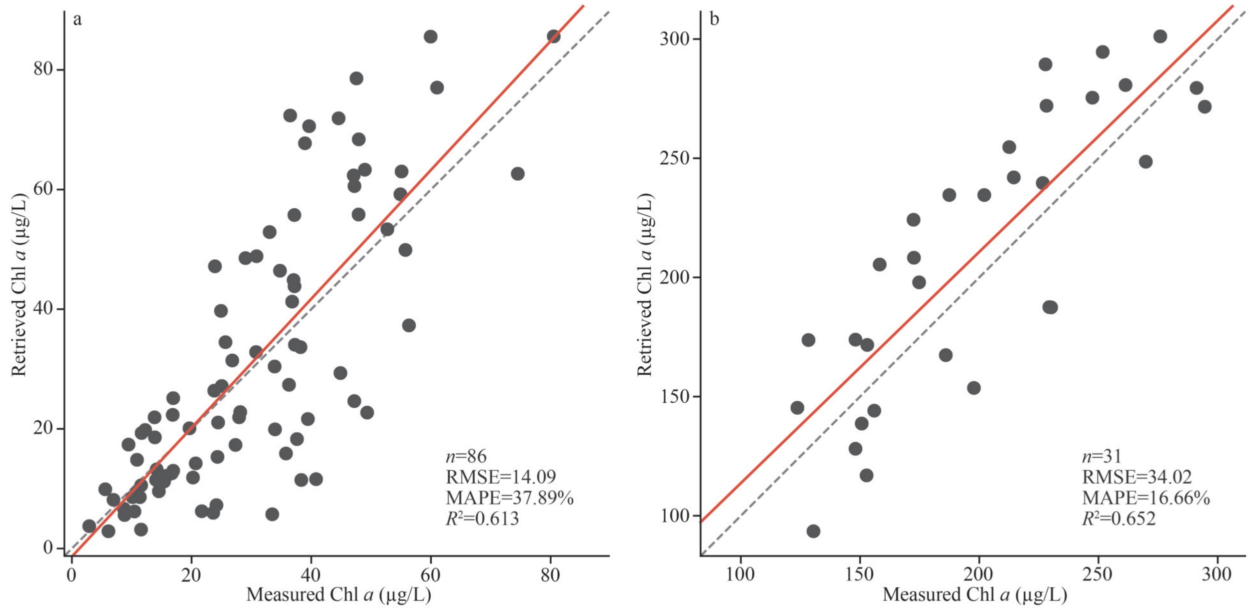

Fig.4 Validation of chlorophyll-*a* retrieval models for non-bloom and bloom regions

a. model validation for non-bloom conditions (n=86), using the Floating Band Absorption (FBA) index-based model. The model effectively captures chlorophyll-*a* variability in low-to-moderate concentration ranges, demonstrating robust performance under non-bloom conditions; b. model performance for bloom conditions (n=31) using the chlorophyll-*a* retrieval model based on the red-edge band ratio (BR), validated against in situ measurements and satellite observations. Statistical metrics, including R^2, RMSE, and MAPE, demonstrate strong agreement between the model and observed data. The red solid line represents the fitted linear regression line between measured and retrieved chlorophyll-*a* concentrations, while the gray dashed line denotes the 1∶1 reference line.

to represent the entire water column. In reality, this assumption often fails, leading to systematic biases in PP estimates.

Given the hydrodynamic and trophic similarities between Taihu Lake and nearby Chaohu Lake, we adapted the phytoplankton vertical distribution model developed by Xue et al. (2015) for Chaohu Lake to refine the chlorophyll vertical profile in Taihu Lake. Both lakes are shallow, eutrophic systems prone to frequent cyanobacterial blooms, predominantly by *Microcystis*, which is the dominant species in both ecosystems. Additionally, the lakes share comparable geographical locations and exhibit similar mixing dynamics. Their shallow depths and nutrient-rich conditions foster analogous biological and hydrodynamic processes. These shared characteristics make Chaohu Lake an appropriate reference for adapting the model to Taihu Lake, ensuring accurate vertical chlorophyll-*a* distribution corrections.

To assess the impact of vertical distribution correction on chlorophyll-*a* estimates, we compared chlorophyll-*a* concentration profiles derived from both the uniform profile assumption and the corrected vertical profile model. Calculations were based on the four profile models outlined in Table 2. By integrating these models across varying depths, we computed the total chlorophyll-*a* concentration for each profile type. Our findings reveal that the uniform profile assumption tends to overestimate chlorophyll-*a* concentrations, particularly under bloom conditions, when chlorophyll is concentrated near the surface. In contrast, the vertical distribution correction yields more accurate estimates by

Table 2 Four types of chlorophyll-*a* vertical profiles and their fitted functions (adapted from Xue et al. (2015))

Class	Chlorophyll-*a* vertical profile class	Fitted function
1	Uniform	$f_1(z)=C$
2	Gaussian	$f_2(z)=22.42+\dfrac{34.08}{0.2\times\sqrt{2\pi}}\times\exp\left[-\dfrac{1}{2}\times\left(\dfrac{z}{0.2}\right)^2\right]$
3	Exponential	$f_3(z)=280.98\times\exp(-3.15\times z)$
4	Hyperbolic	$f_4(z)=29.01\times z^{-0.71}$

accounting for chlorophyll-*a* stratification at different depths, especially during bloom events.

Figure 5 illustrates the decrease in chlorophyll-*a* concentration with increasing depth across all models. The Gaussian (Class 2) and Exponential (Class 3) models show a more rapid decline in chlorophyll-*a* concentration with depth compared to the Hyperbolic (Class 4) model. In contrast, the Uniform profile assumption (Class 1) maintains a constant concentration, leading to an overestimation of chlorophyll-*a*, especially at greater depths. The discrepancy from the Uniform profile is particularly pronounced under bloom conditions, emphasizing the critical role of vertical distribution models in enhancing the accuracy of chlorophyll-*a* estimates.

To determine the appropriate vertical profile class for each observation, a decision-tree classification was used based on bloom status and surface wind speed. The first split criterion was the presence of cyanobacterial blooms (Section 2.3.3.1). In the absence of blooms, vertical profiles are typically Uniform (Class 1) or Gaussian (Class 2), whereas in the presence of blooms, profiles tend to follow Exponential (Class 3) or Hyperbolic (Class 4) shapes.

Surface wind speed was then used to further refine the classification. Under non-bloom conditions, wind speeds > 2.75 m/s favored Class 1 (well-mixed Uniform profile), while calmer conditions (wind ≤2.75 m/s) may have contributed to Class 2 profiles. Under bloom conditions, higher winds (>1.75 m/s) were associated with Class 3 profiles, and lower winds (≤1.75 m/s) with Class 4. Wind data were sourced from ERA5-Land, and wind speed was calculated from the 10-m U and V wind components.

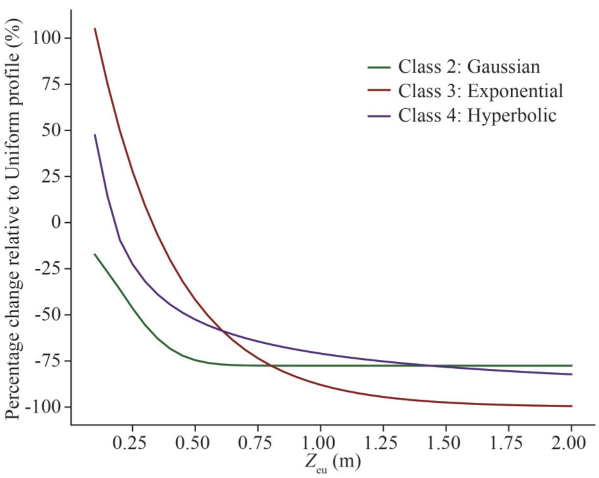

Fig.5 The chlorophyll-*a* concentration profiles, relative to a Uniform distribution (Class 1), are presented across varying depths in Taihu Lake

The percentage change in chlorophyll-*a* concentration is shown for three different distribution models: Gaussian (Class 2), Exponential (Class 3), and Hyperbolic (Class 4). The plot demonstrates how chlorophyll-*a* concentration profiles vary depending on the underlying distribution assumptions. Depth is measured in meters (Z_{eu}), and the percentage change is calculated relative to a uniform surface concentration of 100 μg/L.

Wind-speed thresholds were adopted from Xue et al. (2015), who established a quantitative link between wind forcing and chlorophyll-*a* vertical structure using field observations and a classification and regression tree (CART) framework. Their analysis demonstrated a systematic transition from vertically homogeneous to increasingly stratified phytoplankton distributions with decreasing wind intensity, highlighting wind-driven mixing as a dominant regulator of vertical organization. Optimal thresholds (2.75 and 1.75 m/s) were identified through CART training, cross-validation, and sensitivity analyses that maximized classification accuracy and stability. To ensure transferability under Taihu Lake conditions, we reassessed the robustness of these thresholds using local observational datasets. Analyses confirmed a consistent progression from well-mixed to stratified water-column states with declining wind speed, and sensitivity tests showed stable classification performance around the adopted thresholds. Together, these results support wind forcing as a physically interpretable and statistically robust control on phytoplankton vertical structure in shallow eutrophic lakes.

2.3.5 Analysis of driving factors

We applied an Extreme Gradient Boosting (XGBoost, XGB) model to quantitatively assess the contributions of various environmental factors to the spatiotemporal variability of PP. XGBoost, a gradient-boosted ensemble of decision trees, effectively captures nonlinear relationships and higher-order interactions between PP and its potential drivers (Chen and Guestrin, 2016). The input variables to the model included wind speed (WS), precipitation (Precip), sunlight duration (SD), WST, NTU, EC, DO, NH_3-N, TN, TP, and PAR. PP and all environmental variables were aggregated to monthly mean values for analysis. In this context, WS and Precip represent climatic factors; NTU (turbidity) and PAR reflect light availability conditions; TN and TP represent nutrient levels; and WST, EC, and DO characterize water quality parameters.

Prior to modeling, predictor variables were examined to avoid multicollinearity. The final predictor set thus included PAR, WS, Precip, WST, TP, TN, NTU, EC, and DO, capturing a broad range of light, climate, nutrient, and water-quality conditions (Fig.6). Notably, we included EC and DO to account for variations in ionic strength, external loadings, and biological activity; including these variables improved model performance while maintaining interpretability.

3 RESULT

3.1 Interannual variation in modeled PP in Taihu Lake

Between 2019 and 2025, satellite-derived primary productivity in Taihu Lake exhibited pronounced interannual variability (Fig.7). The annual mean daily modeled PP ranged from 820.68 to 1 527.11 mg C/(m^2·d), with an overall mean of ~1 232.28 mg C/(m^2·d). Exceptionally high modeled productivity was recorded in 2019 and 2020, reaching 1 500.06 and 1 527.11 mg C/(m^2·d), respectively, indicating a period of persistently elevated phytoplankton carbon fixation. The modeled PP declined to 1 139.54 mg C/(m^2·d) in 2021, followed by a rebound to 1 468.05 mg C/(m^2·d) in 2022, suggesting a transient return to a high-productivity state.

A marked decrease in modeled PP was observed after 2022, with productivity dropping to 820.68 mg C/(m^2·d) in 2023, the lowest value recorded during the observation period. Partial recovery was observed in 2024 (~1 131.73 mg C/(m^2·d)) and 2025 (~1 038.78 mg C/(m^2·d)), although these values remained well below the earlier peaks. Collectively, the interannual trajectory from 2019 to 2025 can be characterized as a sequence of “high stability, decline, recovery, further decline, and gradual restoration”. The pronounced temporal variability in modeled PP suggests a dynamic productivity regime and indicates complex interannual ecological dynamics in the lake during this period.

The interannual variability in modeled PP variability in Taihu Lake are linked predominantly to shifts in the dominant algal community and associated environmental changes. Recent studies report a persistent decline in cyanobacterial bloom frequency (particularly of the genus *Microcystis*) since the early 2020s, with the maximum annual bloom area in 2023 being ~76.9% smaller than that observed in 2016 (Song et al., 2024). The observed decline in satellite-derived PP, following a period of

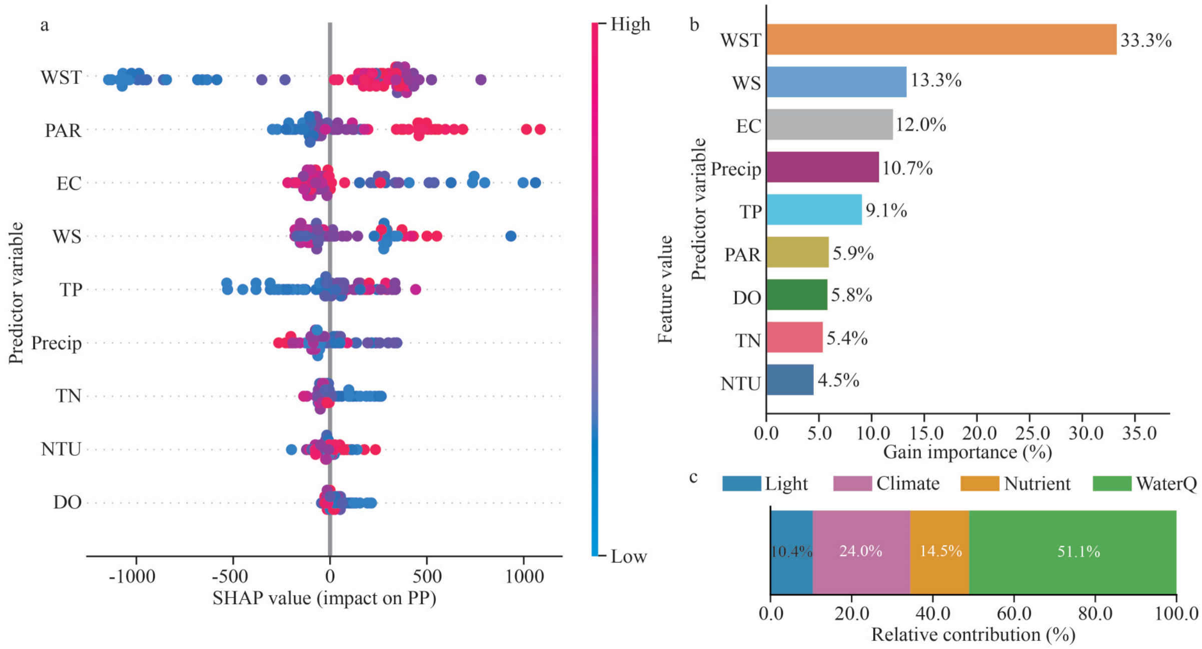

Fig.6 Importance of environmental drivers influencing modeled PP in Taihu Lake (XGBoost results)

a. SHAP summary plot for the test set (chronologically last 20% of data). Each point represents one observation; color indicates the feature value (blue=low, red=high). Positive SHAP values correspond to an increase in predicted PP; b. gain importance of each predictor in the trained XGBoost model, expressed as the percentage of total gain contributed by that feature; c. category-level importance, obtained by summing the gain of individual features within four groups—light availability (PAR, NTU), climate (WS, Precip), nutrients (TP, TN), and water quality (WST, DO, EC)—and normalizing to 100%.

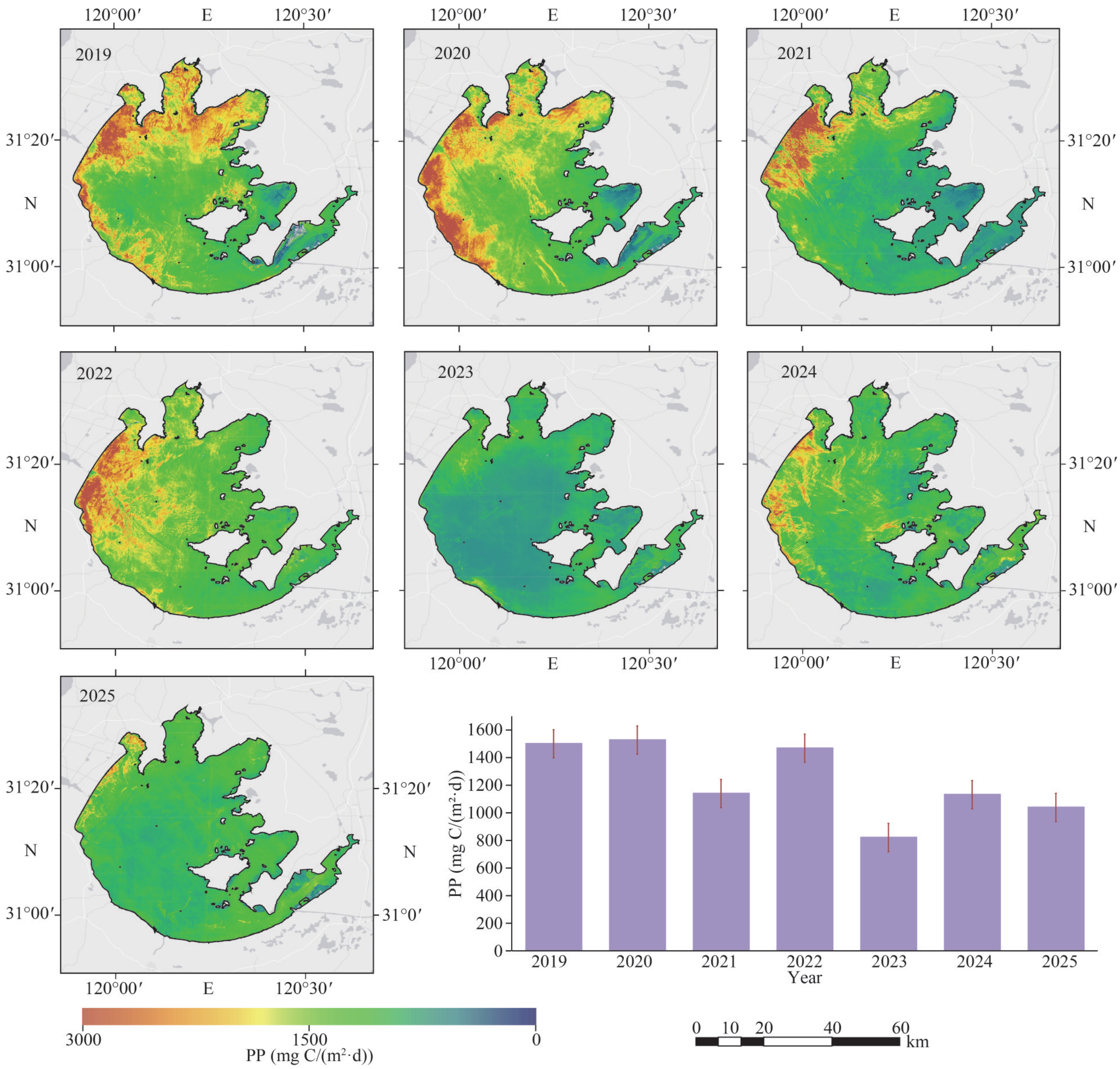

Fig.7 Interannual variation of mean daily satellite-derived PP in Taihu Lake (2019–2025)

Annual mean PP values for each year are shown, illustrating the overall rising and falling trends, as well as the notably low productivity in 2023.

cyanobacterial dominance, may be associated with a reduced efficiency of photosynthetic carbon capture at the system level, potentially contributing to an overall decline in estimated productivity. In addition to this biological restructuring, physical, chemical, and anthropogenic factors have also modulated interannual productivity. Abnormal meteorological conditions in 2023—such as reduced precipitation and weaker winds during the winter–spring transition—diminished external nutrient inputs and suppressed the resuspension and vertical mixing of nutrient-rich bottom sediments. These environmental changes may have reduced nutrient availability in the water column, potentially contributing to lower modeled PP.

Meanwhile, basin-wide nutrient reduction efforts and ecological restoration programs in the watershed have significantly improved water quality over this period; however, they may have temporarily exacerbated nutrient limitation for phytoplankton. The abrupt decline in modeled PP observed in 2023 thus reflects the synergistic effects of cyanobacterial collapse, reduced nutrient supply due to climatic anomalies, and aggressive management interventions. The modest rebound in 2024–2025 suggests that the lake ecosystem is gradually reorganizing toward a new dynamic equilibrium under these altered environmental constraints.

3.2 Monthly dynamic and pattern characteristic

Phytoplankton modeled PP in Taihu Lake showed apparent monthly variations throughout the study period (Fig.8). Modeled PP was lowest in January, followed by a gradual increase beginning in February and an initial peak in May. A moderate decline in June and July preceded a more pronounced peak in August and September. From October onward, modeled PP steadily declined, returning to low levels by November and December. This pattern reveals a characteristic bimodal distribution throughout the year, characterized by an early-year increase, a mid-year dip, and a late-year decline.

The magnitude of this bimodal pattern was substantial. The average maximum monthly modeled PP occurred in September (~1 832.3 mg C/(m^2·d)), while the minimum occurred in January (~318.2 mg C/(m^2·d)), representing nearly a sixfold

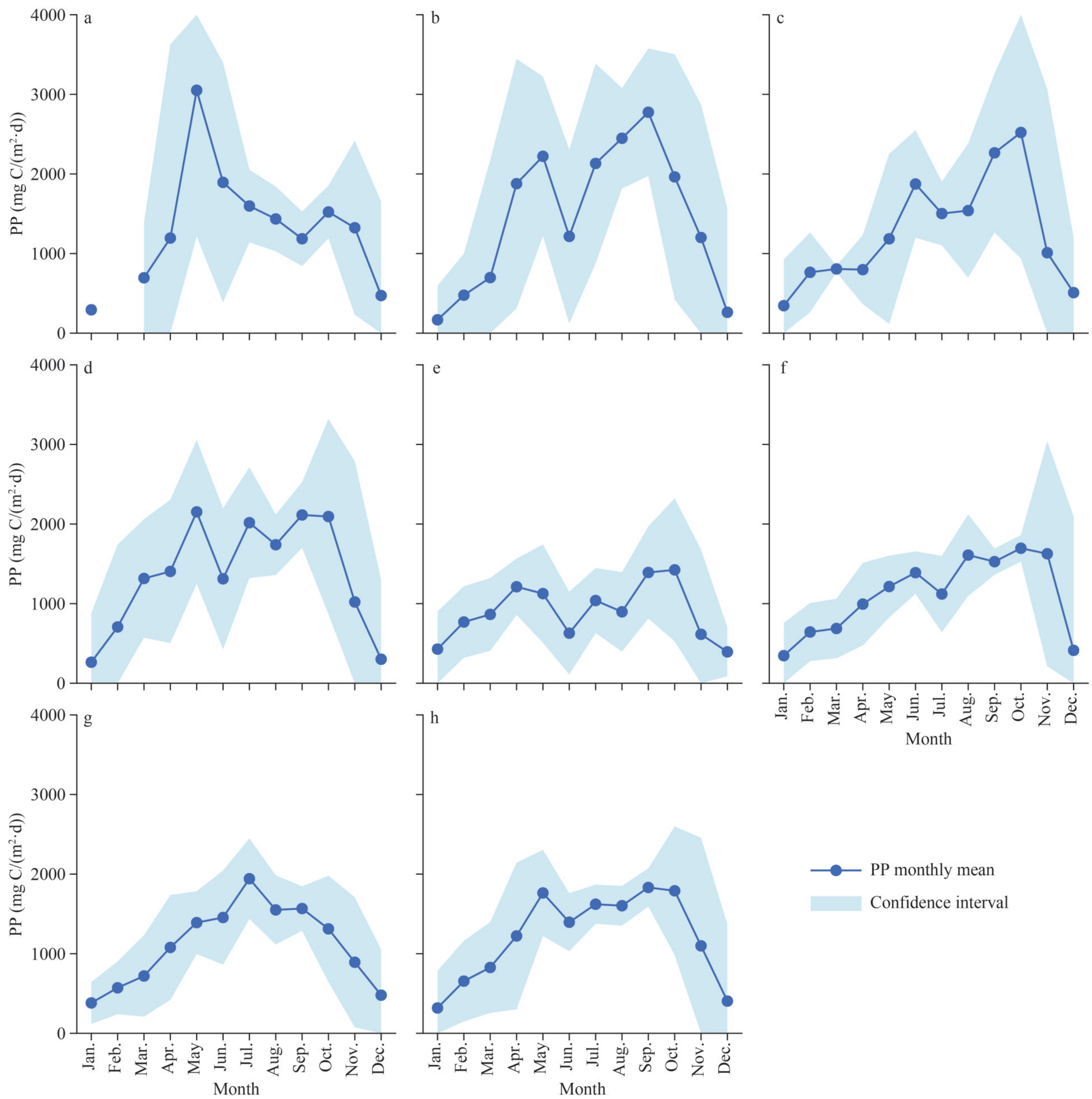

Fig.8 Monthly variations in modeled PP in Taihu Lake (2019–2025)

Panels (a–g) show the monthly mean modeled PP for each year from 2019 through 2025, and panel (h) suggests the climatological monthly mean for the entire period. The bimodal pattern (peaks in late spring and late summer) is evident, alongside interannual differences in peak magnitude.

difference. From July through September, monthly mean modeled PP consistently exceeded 1 600 mg C/(m^2·d), whereas from December through February it remained below 500 mg C/(m^2·d)—a three- to five-fold disparity between warm and cold periods. Notably, September 2021 saw an exceptionally high modeled PP of ~2 265.3 mg C/(m^2·d), substantially higher than values in other years. In contrast, 2024 exhibited relatively muted monthly variability. These findings suggest that interannual climate fluctuations and human activities (e.g., nutrient management) can significantly influence both the timing and magnitude of monthly modeled PP peaks.

Mechanistically, the seasonal rise in modeled PP early in the year can be attributed to increasing solar elevation and day length in March–April, which boost both the intensity and duration of sunlight. This elevates water temperature and stimulates higher maximum carbon fixation rates, creating favorable conditions for phytoplankton growth and a marked increase in modeled PP. By May–June, further warming and improved nutrient availability (after spring mixing) contribute to a secondary peak in productivity. After mid-summer (July onward), although light remains abundant, excessive heating of surface waters begins to inhibit modeled PP; water temperatures often exceed 25 °C in summer, whereas previous studies suggest modeled PP in temperate lakes is optimal around ~19–21 °C. Thermal stress above this optimal range can reduce the photosynthetic efficiency of phytoplankton. Additionally, during mid-summer, the euphotic depth and phytoplankton biomass in Taihu do not increase sufficiently to offset this thermal inhibition, leading to a dip in modeled PP. In October–November, cooling temperatures and autumn mixing promote a brief recovery of chlorophyll concentrations and carbon fixation potential, leading to a small rebound in modeled PP. However, from December into winter, rapidly declining temperatures and light levels, along with lower algal biomass, cause modeled PP to drop sharply to its annual minima and remain low until the following spring.

3.3 Seasonal shift and driving factor

From 2019 to 2025, modeled PP in Taihu Lake showed marked seasonal variation, with peak production in summer and an evident decline through autumn, spring, and winter (Fig.9). Mean daily modeled PP averaged 1 664.41 mg C/(m^2·d) in summer, 1 517.84 mg C/(m^2·d) in autumn, 1 314.39 mg C/(m^2·d) in spring, and only 423.34 mg C/(m^2·d) in winter. This indicates significantly enhanced productivity during the warm seasons compared to the cold season, reflecting the strong temperature dependence of phytoplankton activity.

The frequency distribution of modeled PP across the lake's surface revealed distinct spatial patterns in each season, as indicated by the areal proportions of different productivity ranges. In summer and autumn, virtually the entire lake was in a high-productivity state: areas with >1 000 mg C/(m^2·d) covered ~98.0% (summer) and 93.3% (autumn) of the lake, and extremely high productivity zones (>2 000 mg C/(m^2·d)) accounted for 17.24% and 12.89%, respectively. This suggests that elevated temperatures, strong solar radiation, and intensified sediment nutrient resuspension in summer and autumn jointly enhanced phytoplankton photosynthesis, thereby expanding the spatial extent of high modeled PP areas in Taihu Lake.

In contrast, during winter, most of the lake (≈92.59% of surface area) exhibited very low productivity (0–500 mg C/(m^2·d)). This implies that low water temperatures, limited light availability, and stable water column stratification in winter suppress the release of nutrients from sediments and inhibit algal growth. Spring represented a transitional period, with moderate production zones (1 000–1 500 mg C/(m^2·d)) covering approximately 67.84% of the lake, situated between the low productivity of winter and the high productivity of summer/autumn. This pattern indicates that increasing sunlight and rising water temperature in spring gradually enhance phytoplankton activity. Overall, the seasonal and spatial distribution of modeled PP in Taihu Lake reflects the combined influence of climatic forcing, hydrodynamic conditions, and nutrient cycling, emphasizing the high environmental sensitivity and spatial heterogeneity of phytoplankton productivity in this shallow eutrophic lake.

3.4 Spatial distribution characteristic of modeled PP

The multi-year average daily modeled PP in Taihu Lake exhibits pronounced spatial heterogeneity (Fig.10). Overall, productivity declines from the nutrient-rich bays toward the central lake, and from the western shoreline toward the east. Zhushan Bay had the highest mean modeled PP (~1 667.93 mg C/(m^2·d)), followed by the Coastal Zone (~1 492.14 mg C/(m^2·d)) and Meiliang Bay (~1 386.01 mg C/(m^2·d)). The

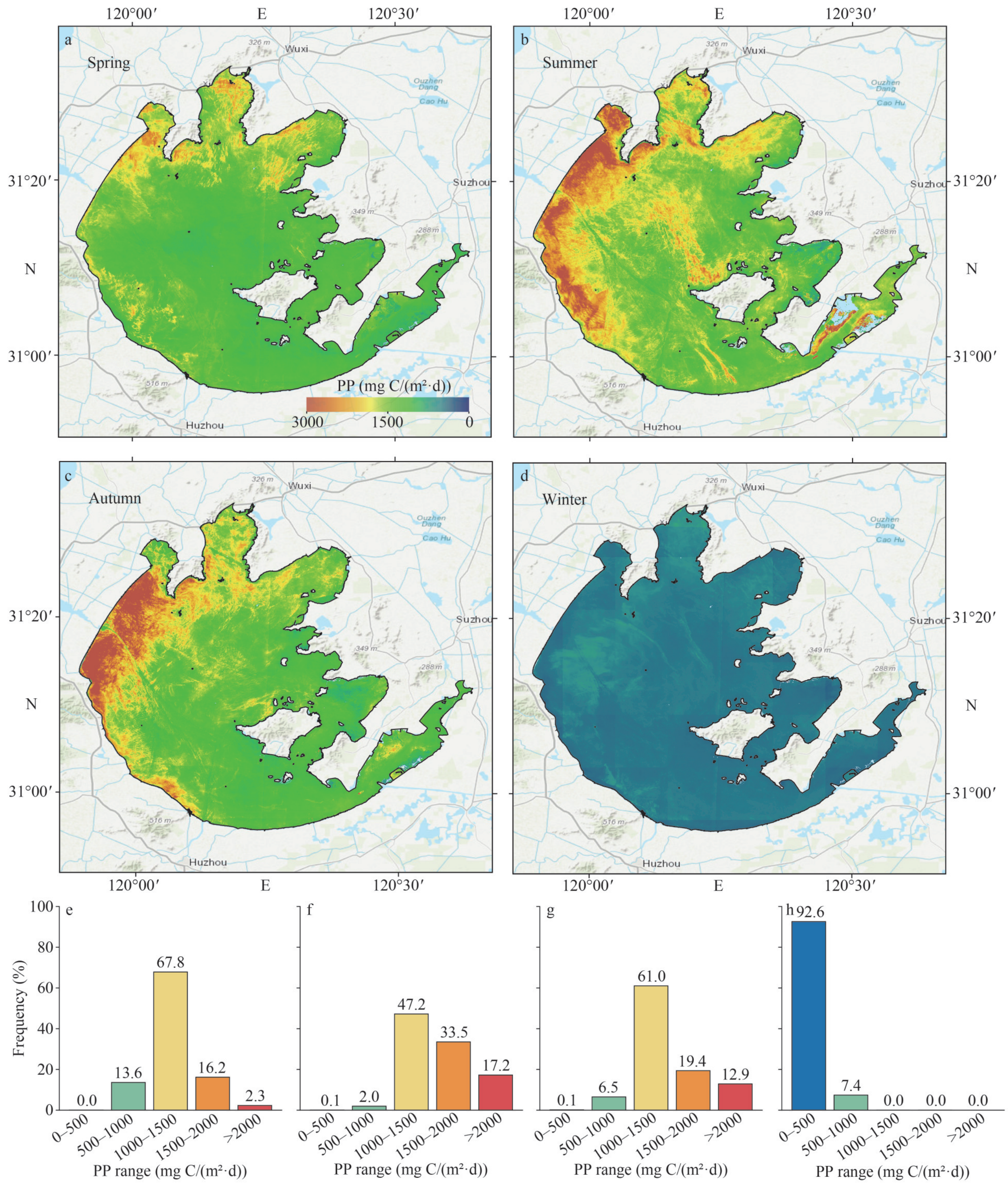

Fig.9 Seasonal distribution of phytoplankton primary productivity in Taihu Lake (2019–2025)

Panels (a–d) show the spatial distribution of mean daily modeled primary productivity (PP) in spring, summer, autumn, and winter, respectively. Warmer colors indicate higher PP values, while cooler colors indicate lower PP values. Panels (e–h) show frequency distributions of mean daily modeled PP for spring, summer, autumn, and winter, respectively. Seasons are defined as spring (March–May), summer (June–August), autumn (September–November), and winter (December–February). Warmer seasons show a broad shift toward higher modeled PP values compared to winter.

Central Lake Area showed intermediate productivity (~1 185.96 mg C/(m^2·d)), whereas eastern regions like East Taihu Lake (~943.24 mg C/(m^2·d)) and Xukou Bay (~969.61 mg C/(m^2·d)) had substantially lower levels.

Coupled hydrological and hydrodynamic processes drive this spatial pattern. Nutrient-enriched runoff from the watershed, particularly into western bays and nearshore areas, delivers high nutrient loads that stimulate phytoplankton growth. Additionally, prevailing wind patterns tend to push surface waters (and floating algae) toward the western and northwestern parts of the lake, causing water convergence and phytoplankton accumulation in those areas. This leads to higher chlorophyll-*a* concentrations and modeled PP in bays like Zhushan and Meiliang. In contrast, the central open waters experience more vigorous vertical mixing and less retention of phytoplankton, limiting productivity. In the east, the extensive presence of submerged and emergent macrophytes likely suppresses phytoplankton productivity by competing for light, nutrients, and space.

Beyond the overall spatial means, modeled PP dynamics showed distinct seasonal behavior in different regions of the lake. Zhushan Bay and Meiliang Bay consistently exhibited elevated PP in spring and summer, with peak daily values exceeding 3 000 mg C/(m^2·d) in some years. These exceptionally high levels reflect sustained external nutrient inputs and favorable conditions in these semi-enclosed bays. In contrast, East Taihu Lake and Xukou Bay experienced much more subdued seasonal fluctuations, with summer peaks typically below 1 500 mg C/(m^2·d). This is likely due to their dense macrophyte cover, which limits phytoplankton growth through shading and nutrient uptake. The Central Lake Area maintained intermediate PP levels with relatively dampened seasonal amplitudes, indicative of moderate nutrient input and continuous mixing.

Across most subregions, modeled PP peaked in spring or summer, began to decline in autumn, and reached annual minima in winter (generally <500 mg C/(m^2·d)), in line with seasonal trends in temperature and solar radiation. Year-to-year variability in modeled PP was most pronounced in the western bays—particularly Zhushan Bay—where spring and summer productivity in years like 2021 and 2025 far exceeded that in other years, pointing to a high sensitivity to interannual climatic and hydrological fluctuations. In contrast, the lake-wide average modeled PP (integrating all regions) tended to mask these extreme regional highs and lows, underscoring the importance of spatially resolved analyses. Overall, the spatiotemporal patterns in Taihu Lake's modeled PP reflect the interplay of nutrient inputs, physical hydrodynamics, and aquatic vegetation distribution, offering insights into the functioning of large, shallow eutrophic lakes.

4 DISCUSSION

4.1 Quantitative uncertainty and sensitivity analysis of modeled PP

4.1.1 Consistency with previously reported PP estimates

As an initial external validation, we compared our PP estimates with previously published values for Taihu Lake and its representative bays. Deng et al. (2017) reported a mean daily PP of 1 094±720 mg C/(m^2·d) during 2003–2013. Yin et al. (2012) documented a mean daily PP of 1 291.61 mg C/(m^2·d) for Meiliang Bay in 2009.

Our lake-wide estimates for 2019–2025 range from approximately 824 to 1 488 mg C/(m^2·d), with a multi-year mean of ~1 256 mg C/(m^2·d). The modeled PP for Meiliang Bay (1 448.47 mg C/(m^2·d)) is slightly higher than the 2009 value but remains within the variability expected for a hyper-eutrophic system.

Although our modeled PP estimates fall within the range of previously reported values for Taihu Lake and its bays, direct comparison should be interpreted with caution. Differences in methodology, spatial and temporal resolution, and model structure (e.g., VGPM parameterization versus other approaches or in situ methods) can introduce systematic discrepancies. Earlier studies were primarily based on different datasets and modeling frameworks, whereas our estimates rely on Sentinel-2 satellite inputs and a customized VGPM implementation. Thus, the apparent agreement supports the general plausibility of our modeled PP estimates but should not be construed as direct validation against in situ productivity measurements.

4.1.2 Structured sensitivity analysis of VGPM inputs

To assess the robustness of PP estimates, we performed a controlled perturbation analysis of the principal VGPM inputs: C_{opt}, Z_{eu}, E_0, and WST, the latter influencing modeled PP through the temperature-dependent term $P^{\mathrm{B}}_{\mathrm{opt}}(T)$.

Each parameter was perturbed within a physically reasonable uncertainty range: C_{opt} (±10%), Z_{eu} (±10%), E_0 (±10%), and WST (±1 K). The relative change in the modeled PP (ΔPP%) was calculated while

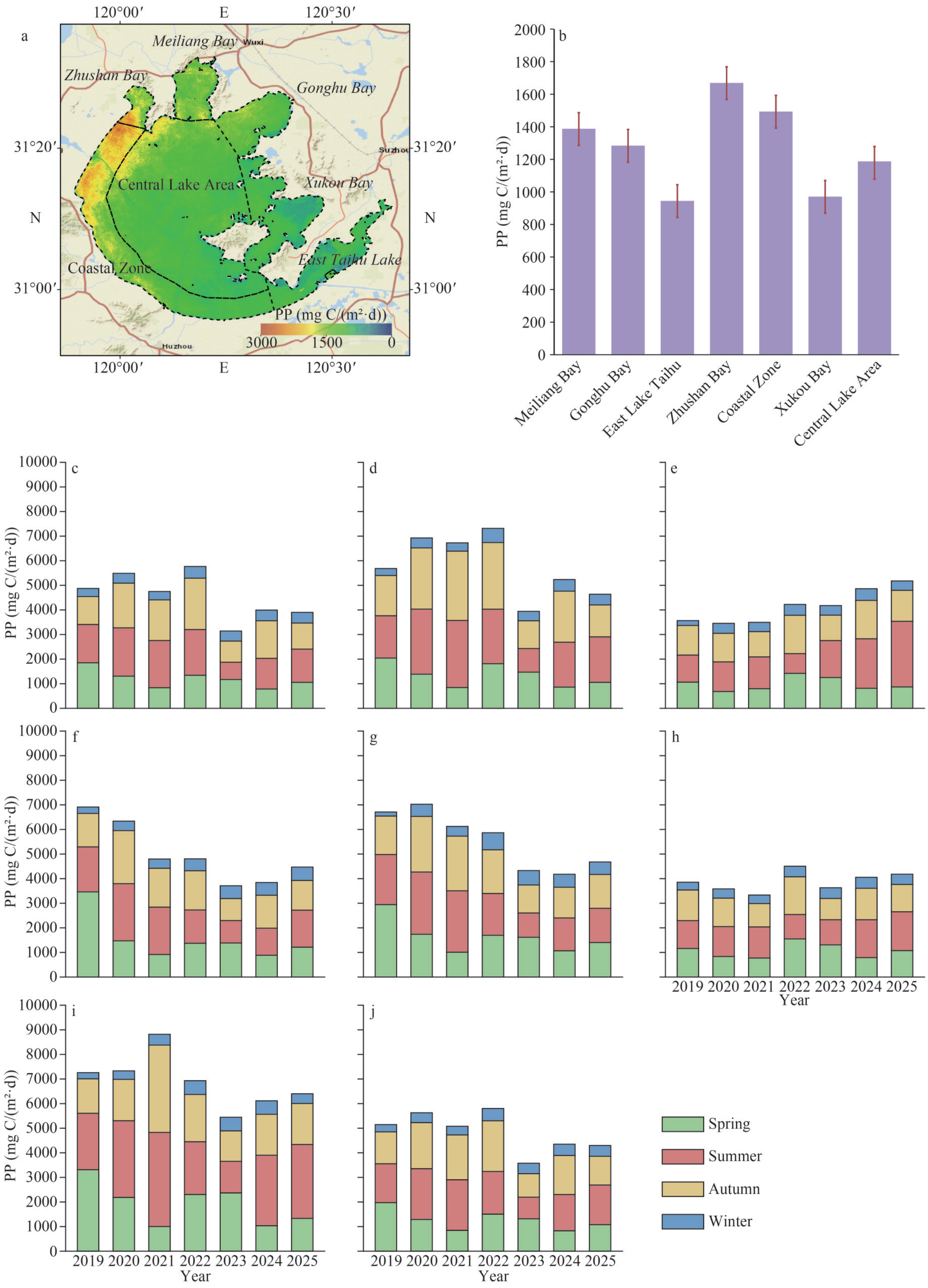

Fig.10 Spatial patterns of modeled PP in Taihu Lake (2019–2025)

a. mean daily modeled PP (mg C/(m²·d)) across Taihu Lake averaged over 2019–2025, illustrating the gradient from high-PP bays (red) to lower modeled PP open waters (blue); b. mean daily modeled PP for each of the seven lake subregions over 2019–2025; c–j. seasonal variation in modeled PP within specific regions: Central Lake, Coastal Zone, East Taihu Lake, Gonghu Bay, Meiliang Bay, Xukou Bay, Zhushan Bay, and the entire lake (respectively).

holding all other variables constant. For E_0 and WST, uncertainty propagation was evaluated across their observed distributions, and the median response and 5th–95th percentile ranges were reported (Fig.11c–d).

The results reveal a clear hierarchy of parameter influence. Modeled PP responds nearly linearly to perturbations in C_{opt} and Z_{eu} (Fig.11a–b); a ±10% change in either parameter produces an approximately proportional ±10% variation in PP. In contrast, PP is less sensitive to E_0, with a median |ΔPP| of 1.9% under ±10% perturbation (Fig.11e). Temperature exerts a moderate, nonlinear influence via $P^B_{opt}(T)$, yielding a median |ΔPP| of 5.8% under ±1 K perturbation. Overall, the summary statistics (Fig.11e) indicate that uncertainties in C_{opt} and Z_{eu} dominate PP variability, whereas radiation and temperature contribute secondary but non-negligible effects.

Despite this general agreement with past studies, several sources of uncertainty affect our PP estimates. First, the phytoplankton vertical distribution model used for correcting chlorophyll profiles was adapted from data collected in Chaohu Lake. Taihu Lake and Chaohu Lake differ in hydrodynamics, optical properties, and algal communities, so some model parameters may not be fully transferable between the lakes. Second, a key VGPM parameter in our model, Z_{eu} was not measured directly but was inferred from empirical temperature-based relationships. Any uncertainty in this parameter could propagate through the PP calculations and introduce error. Third, limitations in data availability (e.g., periods of cloud cover with no usable satellite imagery, such as the entirety of February 2019) reduce the continuity and representativeness of our time series. Gaps or interpolation during such periods can influence the annual means and trends.

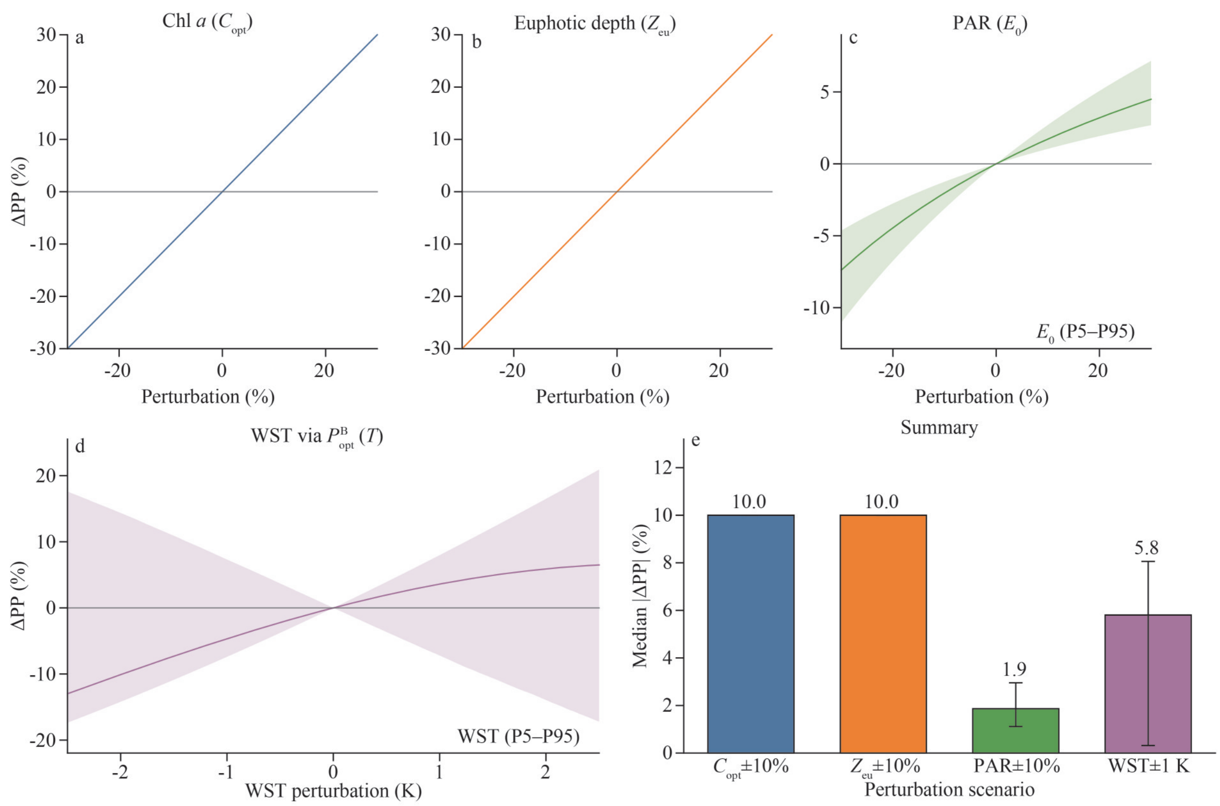

Fig.11 Global sensitivity of VGPM-derived PP to key input parameters

a. relative modeled PP response to perturbations in chlorophyll-*a* concentration (C_{opt}); b. relative PP response to perturbations in euphotic depth (Z_{eu}); c. sensitivity of PP to photosynthetically active radiation (E_0), presented as the median response (solid line) with the 5th–95th percentile envelope (shaded) based on the observed E_0 distribution; d. sensitivity of PP to water surface temperature (WST), propagated through the temperature-dependent $P^B_{opt}(T)$ term, shown as the median response (solid line) and the corresponding 5th–95th percentile envelope (shaded) derived from the observed WST distribution; e. summary of the median absolute change in PP (|ΔPP|) under the prescribed uncertainty ranges (C_{opt}±10%, Z_{eu}±10%, E_0±10%, WST±1 K); error bars denote the 5th–95th percentile interval.

4.2 Analysis of the driving factors of modeled PP

To gain deeper insight into the drivers of modeled PP variability, we implemented a time-aware machine learning analysis. To prevent temporal leakage, we trained and tested the XGBoost model using a chronological split, with the training set consisting of 60 monthly-aggregated samples from 2019 to 2023 and the test set comprising 24 monthly-aggregated samples from 2024 to 2025. The model performance is reported based on the test set, with the following results: RMSE=758.54, mean absolute error (MAE)= 504.21, and R^2=0.423. The feature importance results, interpreted via both gain-based importance and SHAP values (Fig.6), provide a consistent picture of which factors most strongly influence modeled PP.

Before modeling, we addressed multicollinearity among predictors. The final model included nine predictors spanning light availability, climate, nutrients, and water quality, ensuring a comprehensive assessment of potential drivers. Notably, variables such as PAR and temperature are embedded within the VGPM model structure, which may introduce structural dependencies when interpreting feature importance. In this analysis, we emphasize the "correlation" between predictors and modeled PP, rather than attempting to draw causal inferences.

We also recognize the limitations of aggregating data at the monthly scale. While this approach captures long-term trends, it may overlook short-term dynamics, such as those driven by rapid environmental changes or specific events like algal blooms. As a result, this temporal aggregation may restrict our ability to fully capture the short-term drivers of PP variability.

The XGBoost results indicate that water column conditions were the most influential predictors at the individual variable level. In particular, modeled PP increased with WST up to an optimal range, then leveled off or declined at higher temperatures, suggesting there is a thermal optimum for productivity beyond which additional warming confers little benefit or even harm. The interaction between temperature and water quality factors such as EC and DO underscores the complex interplay of environmental factors influencing PP. EC showed a positive correlation with PP, likely because higher EC in Taihu often reflects greater input of mineral-rich or nutrient-laden inflows (or sediment resuspension), which can support higher algal growth. DO had a complex relationship: under high-PP conditions, the effects of DO were weakly negative, possibly because intense algal production coupled with calm, stratified periods can lead to localized oxygen depletion at night (due to community respiration) or indicate stronger stratification during blooms. These patterns suggest that water-quality indicators may serve as proxies for various physical and biogeochemical processes affecting PP (e.g., EC as a proxy for nutrient loading, DO for community respiration and mixing). Furthermore, the interaction between water quality and climatic factors can amplify or mitigate each other's effects, particularly when combined with environmental stressors like nutrient over-enrichment.

Climatic factors ranked next in importance. Wind speed (WS) exhibited a nonlinear relationship with modeled PP: moderate winds are beneficial as they mix the water column and redistribute nutrients, whereas very strong winds tend to resuspend sediments and increase turbidity, reducing light for photosynthesis. Similarly, precipitation had a dual effect: moderate rainfall delivers nutrients from the watershed that can stimulate algal growth, but heavy rainfall can flush the system or increase turbidity (via runoff and sediment disturbance), thus impairing light conditions. The interplay between climate and water quality, such as the synergistic effect of wind and turbidity on light availability, further complicates PP predictions. Moderate winds promote nutrient distribution, while stronger winds hinder photosynthesis by increasing turbidity, creating a negative feedback loop.

Modeled PP increased with nutrient concentrations (TN, TP) up to a threshold, beyond which the response tended to plateau. In a persistently eutrophic lake like Taihu, phosphorus and nitrogen are abundant; our results suggest that once a moderate nutrient level is reached, additional nutrient input does not translate into higher PP because another factor (light or temperature or mixing) becomes limiting. In other words, modeled PP responded strongly to nutrient increases at the low end of observed TN/TP ranges, but the response flattened out at high nutrient levels, indicating a transition from nutrient limitation to light or other limitations. These interactions between nutrient levels and other environmental drivers highlight the complexity of managing eutrophic lakes, where nutrients often become secondary to other limiting factors such as light availability or thermal stratification once a threshold is exceeded.

When grouping factors into broader categories (Fig.6c), water quality-related variables (WST, EC, DO) collectively explained the largest portion of variability, followed by climate (wind and precipitation), then light availability (PAR, turbidity), and then nutrients. This does not contradict the well-established notion that light availability is the proximate limiting factor for primary production in Taihu's turbid waters. Instead, it highlights that some water quality metrics (like turbidity, which is directly linked to light, and EC, which may correlate with suspended solids) inherently capture light availability effects, and that biological feedbacks (e.g., DO consumption during blooms) are part of the PP dynamics. The analysis of interactions between water quality, climate, and nutrients reveals that no single factor dominates PP variability. Rather, a dynamic and interactive relationship exists, where changes in one factor can amplify or dampen the effects of others. This interconnectedness is critical for understanding the mechanisms driving productivity in such a complex and stressed aquatic ecosystem.

Overall, the variability of PP in Taihu Lake emerges from an interplay between chronic eutrophication, meteorological forcing, and light conditions. Nutrient levels remain high enough year-round that they rarely limit growth (especially phosphorus, given long-term eutrophication), whereas weather events (such as wind storms, droughts, or heavy rains) episodically alter mixing and turbidity, thus controlling the effective light climate for phytoplankton. The interaction between long-term nutrient loading and short-term climatic events complicates PP variability predictions, emphasizing the need to consider both chronic and episodic factors in management strategies. Effective management of the lake should therefore pair long-term nutrient load reduction with measures to mitigate short-term light limitation. This could include close monitoring of meteorological events and turbidity, as well as operational responses such as mixing aerators or other interventions during critical periods, to maintain favorable conditions for a balanced level of productivity and ecological functioning.

4.3 Management implication and policy recommendation

By quantitatively characterizing the spatiotemporal variability of modeled PP in Taihu Lake and pinpointing its key environmental drivers, this study provides scientific guidance for lake management and policy. Our results demonstrate that multiple interacting processes jointly modulate lake primary productivity and are highly sensitive to both climatic fluctuations and anthropogenic pressures. Recognizing this complexity is important for developing adaptive management strategies under changing environmental conditions.

From a practical management perspective, phytoplankton modeled PP serves as a critical indicator of both ecological health and carbon sequestration potential. Sustaining a balanced level of modeled PP (neither too low to undermine the food web, nor too high to trigger severe blooms) is desirable. To achieve this, nutrient inputs must be controlled at the watershed scale. Strengthening integrated watershed management to reduce external nitrogen and phosphorus loading will help mitigate eutrophication at its source. This includes managing agricultural runoff, upgrading wastewater treatment facilities, and restoring wetlands or buffer zones that can help retain nutrients.

Meanwhile, extreme weather events (storms, droughts, heatwaves) should be incorporated into the lake's management plans. The establishment of integrated monitoring networks that combine meteorological, hydrological, and ecological data can facilitate real-time assessments and early warnings of ecosystem responses to events such as heavy rainfall or strong winds. For instance, an abrupt calm and hot period might warrant heightened alert for stratification and low DO in bottom waters, whereas a major storm might prompt monitoring for sediment resuspension effects on water quality. Furthermore, scenario-based management frameworks, underpinned by predictive modeling, offer valuable foresight into the potential outcomes of extreme weather. These tools can guide preemptive actions, helping to mitigate disturbances and enhance the resilience of the ecosystem.

Habitat management measures could also support more stable productivity. Restoring submerged aquatic vegetation in suitable eastern shallow zones can improve water clarity by stabilizing sediments and up taking nutrients, albeit with the trade-off of competing with phytoplankton for those resources. Improved water transparency (through vegetation and sediment management) would reduce light limitation and could shift the lake to a state where phytoplankton production is more top-down controlled (by grazers) rather than bottom-up limited (by light). Additionally, maintaining or creating ecological buffer zones around the lake can filter nutrient inflows and provide refuges or additional habitats that enhance the overall resilience of the lake's ecosystem.

At a broader scale, our findings underscore that lakes like Taihu are significant components of regional carbon budgets. Large, productive lakes have the capacity to absorb CO_2 via photosynthesis and sequester carbon in sediments, potentially offsetting some carbon emissions. Thus, inland waters should be integrated into regional and national carbon accounting and climate mitigation strategies. For China's national goals of carbon peaking and carbon neutrality, protecting and managing large lake ecosystems is an opportunity to leverage natural carbon sinks. This could translate into policy support for lake conservation as a form of carbon offset or eco-compensation, in addition to the traditional motivations of biodiversity and water quality.

In summary, maintaining Taihu Lake's ecological balance will require a multi-faceted approach: sustained nutrient load reductions, adaptation to climate variability (through monitoring and responsive management), and habitat manipulations that improve light conditions and ecosystem resilience. Combined with continued research and monitoring, these strategies will help ensure that Taihu can serve as both a healthy ecosystem and a functional component of the regional carbon cycle under future conditions. Achieving this goal requires a comprehensive lake management framework that integrates climate risk assessments, ecological restoration, and carbon management strategies to ensure the long-term sustainability and resilience of the lake's ecosystem.

4.4 Perspective

This study developed a remote sensing-based framework for estimating PP in Taihu Lake, tailored to the lake's distinct environmental characteristics. The framework integrates satellite retrievals of chlorophyll *a* and euphotic depth with a correction for phytoplankton vertical distribution, optimizing a VGPM model for shallow turbid conditions. These advancements provide a robust methodological foundation for improving our understanding of lake carbon cycling and phytoplankton productivity dynamics.

Nonetheless, our current analysis covers a relatively short time span (2019–2025). The long-term trends and the full range of response mechanisms of PP in Taihu Lake remain insufficiently captured. A priority for future work is therefore to continue monitoring modeled PP in Taihu Lake on a long-term basis. Extending the observation period will help distinguish enduring shifts (e.g., due to climate change or sustained nutrient management) from short-term fluctuations.

As a first step toward continuous monitoring, we have implemented an automated PP estimation algorithm on the GEE platform (accessible at the provided link). This cloud-based system enables real-time data updates and storage, allowing for ongoing spatiotemporal analysis with minimal manual effort (see Code Availability). The operational workflow established here can serve as a reliable tool for evaluating lake carbon sink function and informing management decisions on an ongoing basis.

Looking ahead, future research should advance in several directions. First, there is a need to incorporate multi-source and multi-sensor remote sensing data to complement Sentinel-2. Combining observations from sensors with different strengths (for example, higher-frequency but lower-resolution satellite data, thermal infrared data for temperature, or LiDAR for high-resolution bathymetry) could provide a more comprehensive picture of factors influencing modeled PP. This multi-source integration, along with in situ measurements (e.g., high-frequency buoy data for chlorophyll fluorescence or PAR profiles), will enable more accurate and frequent assessments of lake productivity and its drivers.

Second, further refinement of the model structure is warranted. Machine learning and data assimilation techniques offer promising avenues for improving the retrieval of intermediate parameters, such as euphotic depth or chlorophyll concentration. For instance, using recurrent neural networks or hybrid models could better account for time-lagged effects (e.g., nutrient upwelling after a wind event leading to a bloom days later). Data assimilation, where satellite observations are continuously integrated into a dynamic lake ecosystem model, could also enhance predictive capabilities. It will be important to validate and potentially recalibrate the framework in other lakes (both within China and globally) to ensure its transferability. By comparing results from Taihu Lake with those from lakes of different sizes, depths, and trophic states, we can generalize the approach for regional or national lake carbon assessments.

The findings of this study offer substantial practical applications beyond basic scientific inquiry. The near-real-time monitoring of PP enabled by our GEE-based system can be seamlessly integrated into operational lake management frameworks, strengthening early warning systems and decision-making processes. By combining continuous modeled PP estimates with nutrient loading data and meteorological forecasts, it is possible to predict periods of elevated algal productivity, thereby facilitating the implementation of proactive mitigation strategies, such as targeted nutrient reduction or adaptive water management. Furthermore, the PP products can contribute to regional-scale carbon budget assessments, providing stakeholders and policymakers with actionable metrics to evaluate the effectiveness of water quality improvement initiatives and quantify ecosystem services associated with carbon sequestration.

Ultimately, it is essential to enhance the integration of modeled PP remote sensing data with ecosystem management and decision-making processes. The near real-time modeled PP monitoring enabled by our GEE-based system can be linked with management actions—such as issuing alerts for unusual productivity drops or spikes that might signal developing problems (e.g., a potential fish kill risk or an upcoming algal bloom). Combining our modeled PP dataset with other data streams (nutrient input records, meteorological forecasts, hydrodynamic models) could support the development of intelligent management tools. For example, a decision support system could recommend preemptive measures (like temporary aeration in isolated bays or strategic water diversion) when models predict that certain weather conditions will likely lead to critically low or high modeled PP events. Such proactive management, informed by continuous remote sensing, will enhance the adaptive capacity of lake management in response to both climatic variability and long-term change. An integrated decision support system could automate risk assessment by combining PP anomaly detection with predicted weather events and nutrient inputs, recommending appropriate actions to mitigate adverse events. The operationalization of these research findings will strengthen the adaptive capacity of lake management agencies and support sustainable freshwater resource governance amid climatic and anthropogenic pressures.

5 CONCLUSION

Using high-resolution Sentinel-2 imagery and a tailored modeling approach, this study quantified the spatiotemporal dynamics of phytoplankton primary productivity in Taihu Lake from 2019 to 2025. The results reveal a distinct spatial pattern: modeled PP generally decreases from the eutrophic bays toward the central basin and from west to east across the lake. Zhushan Bay and Meiliang Bay consistently exhibited the highest modeled PP levels, whereas the eastern region of the lake, dominated by dense aquatic vegetation, showed relatively low phytoplankton productivity.

Temporally, modeled PP demonstrated a clear seasonal cycle, with higher values in summer and autumn and lower values in winter and spring, resulting in a characteristic bimodal pattern each year. Elevated modeled PP occurred primarily during periods of warm water temperature and ample sunlight, while increased turbidity (light limitation) suppressed productivity. This indicates that light availability remains a critical constraint on primary production in this shallow, turbid lake. Overall, water temperature, light conditions, and hydrodynamic processes (mixing and stratification) jointly governed the observed spatiotemporal variability of modeled PP, whereas the influence of nutrient enrichment was less apparent under persistently eutrophic conditions.

Methodologically, the incorporation of a phytoplankton vertical distribution correction, separate chlorophyll models for bloom vs. non-bloom waters, and spectral exclusion of macrophyte signals significantly improved the VGPM model's accuracy and applicability in a shallow, highly turbid system like Taihu. The implementation of an automated inversion workflow on the GEE platform enabled efficient estimation and continuous monitoring of lake-wide PP over multiple years. Taken together, this study provides a transferable framework for quantifying carbon fixation in inland waters and offers valuable insights to support the management of eutrophic lakes, the evaluation of carbon sinks, and assessments of ecosystem responses to climate change.

6 DATA AVAILABILITY STATEMENT

Data availability: The Sentinel-2 and ERA5-Land reanalysis datasets used in this study are accessible through the Google Earth Engine data archive (see GEE datasets).

Code availability: The code for the automatic PP calculation algorithm is available on the Google Earth Engine platform (https://code.earthengine.google.co.in/af73ccc19292614b8d3fb7a485082979). Code for other analyses in this study is available from the corresponding author upon reasonable request.

7 ACKNOWLEDGMENT

Data produced and analyzed in this paper were generated in collaboration with the Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences.

References

Anderson S I, Barton A D, Clayton S et al. 2021. Marine phytoplankton functional types exhibit diverse responses to thermal change. *Nature Communications*, **12**(1): 6413, https://doi.org/10.1038/s41467-021-26651-8.

Bachmann R W, Hoyer M V, Canfield D E Jr. 2000. The potential for wave disturbance in shallow Florida lakes. *Lake and Reservoir Management*, **16**(4): 281-291, https://doi.org/10.1080/07438140009354236.

Behrenfeld M J, Falkowski P G. 1997. Photosynthetic rates derived from satellite-based chlorophyll concentration. *Limnology and Oceanography*, **42**(1): 1-20, https://doi.org/10.4319/lo.1997.42.1.0001.

Cassar N, DiFiore P J, Barnett B A et al. 2011. The influence of iron and light on net community production in the Subantarctic and Polar Frontal Zones. *Biogeosciences*, **8**(2): 227-237, https://doi.org/10.5194/bg-8-227-2011.

Chen T Q, Guestrin C. 2016. XGBoost: a scalable tree boosting system. *In*: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Association for Computing Machinery, San Francisco. p.785-794, https://doi.org/10.1145/2939672.2939785.

Cole J J, Prairie Y T, Caraco N F et al. 2007. Plumbing the global carbon cycle: integrating inland waters into the terrestrial carbon budget. *Ecosystems*, **10**(1): 172-185, https://doi.org/10.1007/s10021-006-9013-8.

Deng Y B, Zhang Y L, Li D P et al. 2017. Temporal and spatial dynamics of phytoplankton primary production in Lake Taihu derived from MODIS data. *Remote Sensing*, **9**(3): 195, https://doi.org/10.3390/rs9030195.

Gao Y, Jia J J, Lu Y et al. 2021. Determining dominating control mechanisms of inland water carbon cycling processes and associated gross primary productivity on regional and global scales. *Earth-Science Reviews*, **213**: 103497, https://doi.org/10.1016/j.earscirev.2020.103497.

Gregg W W, Conkright M E, Ginoux P et al. 2003. Ocean primary production and climate: global decadal changes. *Geophysical Research Letters*, **30**(15): 1809, https://doi.org/10.1029/2003GL016889.

Halsey K H, O'Malley R T, Graff J R et al. 2013. A common partitioning strategy for photosynthetic products in evolutionarily distinct phytoplankton species. *New Phytologist*, **198**(4): 1030-1038, https://doi.org/10.1111/nph.12209.

Hou X J, Liu J Y, Huang H B et al. 2024. Mapping global lake aquatic vegetation dynamics using 10-m resolution satellite observations. *Science Bulletin*, **69**(19): 3115-3126, https://doi.org/10.1016/j.scib.2024.05.009.

Kauer T, Kutser T, Arst H et al. 2015. Modelling primary production in shallow well mixed lakes based on MERIS satellite data. *Remote Sensing of Environment*, **163**: 253-261, https://doi.org/10.1016/j.rse.2015.03.023.

Liu J Y, Hou X J, Huang H B et al. 2025b. Transformation of aquatic vegetation in Chinese lakes from 1988 to 2023. *Remote Sensing of Environment*, **330**: 114999, https://doi.org/10.1016/j.rse.2025.114999.

Liu J Y, Huang H B, Hou X J et al. 2025a. Expansion of aquatic vegetation in northern lakes amplified methane emissions. *Nature Geoscience*, **18**(4): 322-329, https://doi.org/10.1038/s41561-025-01667-7.

Ndong M, Bird D, Nguyen-Quang T et al. 2014. Estimating the risk of cyanobacterial occurrence using an index integrating meteorological factors: Application to drinking water production. *Water research*, **56**: 98-108, https://doi.org/10.1016/j.watres.2014.02.023.

Palmer S C J, Kutser T, Hunter P D. 2015. Remote sensing of inland waters: challenges, progress and future directions. *Remote Sensing of Environment*, **157**: 1-8, https://doi.org/10.1016/j.rse.2014.09.021.

Pérez V, Fernández E, Marañón E et al. 2005. Seasonal and interannual variability of chlorophyll *a* and primary production in the Equatorial Atlantic: *in situ* and remote sensing observations. *Journal of Plankton Research*, **27**(2): 189-197, https://doi.org/10.1093/plankt/fbh159.

Platt T. 1986. Primary production of the ocean water column as a function of surface light intensity: algorithms for remote sensing. *Deep Sea Research Part A. Oceanographic Research Papers*, **33**(2): 149-163, https://doi.org/10.1016/0198-0149(86)90115-9.

Shen M, Duan H T, Cao Z G et al. 2017. Remote sensing estimation algorithm of diffuse attenuation coefficient applicable to different satellite data in Lake Taihu, China. *Journal of Lake Sciences*, **29**(6): 1473-1484, https://doi.org/10.18307/2017.0619. (in Chinese with English abstract)

Shih Y Y, Shiah F K, Lai C C et al. 2021. Comparison of primary production using *in situ* and satellite-derived values at the SEATS Station in the South China Sea. *Frontiers in Marine Science*, **8**: 747763, https://doi.org/10.3389/fmars.2021.747763.

Silsbe G M, Oxborough K, Suggett D J et al. 2015. Toward autonomous measurements of photosynthetic electron transport rates: an evaluation of active fluorescence-based measurements of photochemistry. *Limnology and Oceanography, Methods*, **13**(3): 138-155, https://doi.org/10.1002/lom3.10014.

Song T, Zhang H J, Xu Y J et al. 2024. Cyanobacterial blooms in Lake Taihu: temporal trends and potential drivers. *Science of the Total Environment*, **942**: 173684, https://doi.org/10.1016/j.scitotenv.2024.173684.

Tranvik L J, Downing J A, Cotner J B et al. 2009. Lakes and reservoirs as regulators of carbon cycling and climate. *Limnology and Oceanography*, **54**(6 part 2): 2298-2314, https://doi.org/10.4319/lo.2009.54.6_part_2.2298.

Tsubo M, Walker S. 2005. Relationships between photosynthetically active radiation and clearness index at Bloemfontein, South Africa. *Theoretical and applied climatology*, **80**(1): 17-25, https://doi.org/10.1007/s00704-004-0080-5.

Westberry T K, Silsbe G M, Behrenfeld M J. 2023. Gross and net primary production in the global ocean: an ocean color remote sensing perspective. *Earth-Science Reviews*, **237**: 104322, https://doi.org/10.1016/j.earscirev.2023.104322.

Xue K, Zhang Y C, Duan H T et al. 2015. A remote sensing approach to estimate vertical profile classes of phytoplankton in a eutrophic lake. *Remote Sensing*, **7**(11): 14403-14427, https://doi.org/10.3390/rs71114403.

Ye H B, Chen C Q, Sun Z H et al. 2015. Estimation of the primary productivity in Pearl River Estuary using MODIS data. *Estuaries and Coasts*, **38**(2): 506-518, https://doi.org/10.1007/s12237-014-9830-5.

Yin Y, Zhang Y L, Shi Z Q et al. 2012. Estimation of spatial and seasonal changes in phytoplankton primary production in Meiliang Bay, Lake Taihu, based on the vertically generalized production model and MODIS data. *Acta Ecologica Sinica*, **32**(11): 3528-3537, https://doi.org/10.5846/stxb201105070595. (in Chinese with English abstract)

Zhang M, Straile D, Chen F Z et al. 2018. Dynamics and drivers of phytoplankton richness and composition along productivity gradient. *Science of the Total Environment*, **625**: 275-284, https://doi.org/10.1016/j.scitotenv.2017.12.288.

Zhang Y L, Liu X H, Yin Y et al. 2012. A simple optical model to estimate diffuse attenuation coefficient of photosynthetically active radiation in an extremely turbid lake from surface reflectance. *Optics Express*, **20**(18): 20482-20493, https://doi.org/10.1364/OE.20.020482.