Acta Oceanol. Sin., 2017, Vol. 36, No. 6, P. 112–121
DOI: 10.1007/s13131-017-1047-8
http://www.hyxb.org.cn
E-mail: hyxbe@263.net

Assessing the uncertainties of phytoplankton absorption-based model estimates of marine net primary productivity

TAO Zui[1], MA Sheng[2]*, YANG Xiaofeng[1], WANG Yan[3]

[1] State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100101, China

[2] Beijing North-Star Digital Remote Sensing Technology Co. Ltd., Beijing 100120, China

[3] State Environmental Protection Key Laboratory of Numerical Modeling for Environment Impact Assessment, Beijing 100012, China

Received 8 April 2016; accepted 28 July 2016

Abstract

Satellite-derived phytoplankton pigment absorption (a_{ph}) has been used as a key predictor of phytoplankton photosynthetic efficiency to estimate global ocean net primary production (NPP). In this study, an a_{ph}-based NPP model (AbPM) with four input parameters including the photosynthetically available radiation (PAR), diffuse attenuation at 490 nm ($K_d(490)$), euphotic zone depth (Z_{eu}) and the phytoplankton pigment absorption coefficient (a_{ph}) is compared with the chlorophyll-based model and carbon-based model. It is found that the AbPM has significant advantages on the ocean NPP estimation compared with the chlorophyll-based model and carbon-based model. For example, AbPM greatly outperformed the other two models at most monitoring sites and had the best accuracy, including the smallest values of RMSD and bias for the NPP estimate, and the best correlation between the observations and the modeled NPPs. In order to ensure the robustness of the model, the uncertainty in NPP estimates of the AbPM was assessed using a Monte Carlo simulation. At first, the frequency histograms of simple difference (δ), and logarithmic difference (δ^{LOG}) between model estimates and in situ data confirm that the two input parameters (Z_{eu} and PAR) approximate the Normal Distribution, and another two input parameters (a_{ph} and $K_d(490)$) approximate the logarithmic Normal Distribution. Second, the uncertainty in NPP estimates in the AbPM was assessed by using the Monte Carlo simulation. Here both the PB (percentage bias), defined as the ratio of ΔNPP to the retrieved NPP, and the CV (coefficient of variation), defined as the ratio of the standard deviation to the mean are used to indicate the uncertainty in the NPP brought by input parameter to AbPM model. The uncertainty related to magnitude is denoted by PB and the uncertainty related to scatter range is denoted by CV. Our investigations demonstrate that PB of NPP uncertainty brought by all parameters with an annual mean of 5.5% covered a range of –5%–15% for the global ocean. The PB uncertainty of AbPM model was mainly caused by a_{ph}; the PB of NPP uncertainty brought by a_{ph} had an annual mean of 4.1% for the global ocean. The CV brought by all the parameters with an annual mean of 105% covered a range of 98%–134% for global ocean. For the coastal zone of Antarctica with higher productivity, the PB and CV of NPP uncertainty brought by all parameters had annual means of 7.1% and 121%, respectively, which are significantly larger than those obtained in the global ocean. This study suggests that the NPPs estimated by AbPM model are more accurate than others, but the magnitude and scatter range of NPP errors brought by input parameter to AbPM model could not be neglected, especially in the coastal area with high productivity. So the improving accuracy of satellite retrieval of input parameters should be necessary. The investigation also confirmed that the SST related correction is effective for improving the model accuracy in low temperature condition.

Key words: marine net primary production, phytoplankton pigment absorption, satellite remote sensing, uncertainty analysis, Monte Carlo simulation

Citation: Tao Zui, Ma Sheng, Yang Xiaofeng, Wang Yan. 2017. Assessing the uncertainties of phytoplankton absorption-based model estimates of marine net primary productivity. Acta Oceanologica Sinica, 36(6): 112–121, doi: 10.1007/s13131-017-1047-8

1 Introduction

Estimation of marine net primary productivity (NPP) is an important component of global biogeochemical cycling analyses. Nearly half of the global photosynthetically fixed carbon is derived from ocean phytoplankton (Behrenfeld et al., 2001; Falkowski et al., 1998; Milutinović et al., 2009). The NPP has already been studied for many years, and it is closely related to the study of global climate processes (Bianchi et al., 2005). NPP is traditionally measured by ships, but the geographic and temporal coverage of a ship is limited by the high cost of shipboard measurements (Milutinović and Bertino, 2011; Carr et al., 2006; Saba et al., 2010). Fortunately, with the development of satellite technology, inversion methods for NPP based on remote sensing data were gradually developed (Saba et al., 2010; McClain et al., 1998). These methods would greatly improve the efficiency of NPP estimation (Friedrichs et al., 2009).

Foundation item: The National Natural Science Fundation of China under contract No. 41501389; the Foundation of State Key Laboratory of Remote Sensing Science in China under contract No. OFSLRSS201509.
*Corresponding author, E-mail: masheng_1987@163.com

Many satellite-based NPP models have been previously proposed (Saba et al., 2011; Friedrichs et al., 2009; Siegel et al., 2001; Antoine and Morel, 1996; Longhurst et al., 1995; Ondrusek et al., 2001; Platt and Sathyendranath, 1993; Westberry et al., 2008; Behrenfeld and Falkowski, 1997). Most of these models are based on the chlorophyll *a* concentration (Chl) and the phytoplankton carbon biomass (C). Recently, many field measurements have demonstrated that variations in productivity are more closely related to phytoplankton absorption than to chlorophyll *a* concentration (John Marra et al., 2003; Medina-Gómez and Herrera-Silveira, 2006; Muller-Karger et al., 2001). Researchers have also indicated that the absorption of phytoplankton can be used to build more reasonable NPP models (Lee et al., 1996; Hirawake et al., 2011). Thus, an a_{ph}-based NPP model (AbPM) was rapidly developed and successfully applied to the regional ocean NPP calculations (Ma et al., 2014; Lee et al., 2011).

However, several problems still exist for the global use of the AbPM model (Ma et al., 2014). Net primary productivity fields, derived from satellite observations of ocean color, should be commonly published with relevant information on uncertainties. Many studies have investigated the precision of the Chl~ and C~ NPP model and the uncertainty of the input parameters (Saba et al., 2010; Friedrichs et al., 2009). Unfortunately, there is insufficient research on these issues for the AbPM model. According to our research, the AbPM model may possibly overestimate NPP in the coastal zone areas (especially for the coastal zone around Antarctica). The possible reason for this phenomenon is that properties of phytoplankton may be more complex in these areas. Some studies also indicate that phytoplankton cell size, iron content, nutrients and the sea surface temperature may affect the environmental adaptability of phytoplankton, which would in turn affect the a_{ph} based model's stability (Sunda and Huntsman, 1997; Kiefer and Cullen, 1991; Marañón et al., 2012). Nevertheless, these possible reasons should to be examined further.

In this study, we investigated the factors which may affect the accuracy of the AbPM model. Comparisons and validation of model outputs versus *in situ* data were performed for the AbPM model and other two models. The Monte Carlo simulation was used to assess the uncertainties of input parameters of the AbPM model. The overall uncertainty and individual uncertainty of the input parameters were evaluated and the impact of the input parameters on the AbPM model was subsequently discussed. Additionally, the influence of the sea surface temperature (SST) on the AbPM model was also be considered. A temperature function was used to revise the limitation of low temperature in the model. Then quantitative analysis was used to detect the causes of errors in the accuracy of the AbPM model, which would be helpful for the improvement of the model.

2 Data and methods

2.1 *Remote sensing data*

Moderate Resolution Imaging Spectroradiometer (MODIS) products were collected as the input datasets, which included photosynthetically available radiation (PAR), diffuse attenuation at 490 nm (K_{d} (490)), euphotic zone depth (Z_{eu}) and the phytoplankton pigment absorption coefficient ($a_{\mathrm{ph}}(\lambda)$). All of these products covered the period from January 2003 to December 2012, and the spatial resolutions were re-sampled to 9 km. In these products, monthly averaged data were used to calculate the monthly averaged NPP, and daily data synchronized with *in situ* data were used for validation. Nitracline depths (Z_{no3}) were calculated from monthly climatological nutrient data obtained from the World Ocean Atlas and defined as the depth where nitrate + nitrite first exceeded 0.5 μmol/L (Garcia, 2010).

2.2 *In situ data*

Various NPP *in situ* data were collected from different projects during the period between 2003 and 2012. After matching with satellite-derived data, 773 NPP sample sites over six regions (HOT, BATS, CALCOFI, DYFAMED, AMLR and WAP) were finally chosen (Table 1). These *in situ* data were measured based on dawn-to-dark *in situ* ^{14}C incubations from the surface to the 1% light-level at various depth intervals. At some time-series monitoring sites (such as HOT, BATS and CALCOFI), NPP measurements were generally conducted monthly (HOT and BATS) or quarterly (CALCOFI). At several ship monitoring sites (DYFAMED, AMLR and WAP), NPP was measured during navigation. Figure 1 shows the locations of all the NPP *in situ* measurement stations. The *in situ* measurements for four parameters ($a_{\mathrm{ph}}(\lambda)$, *PAR*, K_{d} (490) and Z_{eu}) were collected from the website of SeaWiFS Bio-optical Archive and Storage System (SeaBASS). Samples of $a_{\mathrm{ph}}(\lambda)$ over the 400–700 nm spectral region were measured in collected water samples using the GF/F filter-pad transmission technique (Marra et al., 2007). In this article, *in situ* data of four parameters were only used for error detection and analysis of the MODIS products. Figure 2 shows the locations of *in situ* measurement stations for all input parameters.

2.3 *AbPM model*

The availability of operational a_{ph} production and field investigations has encouraged investigators to apply satellite-derived data to estimate NPP on a global oceanic scale. The following is a spectrally integrated formulation for AbPM at depth z (Ma et al., 2014):

$$NPP(z) = \phi_{\max} \frac{K_{\mathrm{m}}}{K_{\mathrm{m}} + E(z)} \left[\bar{a}_{\mathrm{ph}}(z) \times E(z) \times \exp\left(-\nu \times E(z)\right)\right], \quad (1)$$

where *NPP* is net primary productivity, z is depth, and all the parameters are listed in Table 2. In Eq. (1), $E(z)$ is the integration of downwelling solar irradiance, which can be calculated as follows (Ricchiazzi et al., 1998):

Table 1. Description of each region which was selected as the NPP field data

General region	Program	Ecosystem type	Count	Sampling time range (year/month)	Spatial coverage	NPP method
Northeast Atlantic Ocean	BATS	subtropical gyre	91	2003/01–2010/12	single station	*in situ*
North Pacific Ocean	HOT	subtropical zone	78	2003/01–2010/12	single station	*in situ*
California Coast	CALCOFI	temperate zone	38	2003/01–2012/03	single station	*in situ*
Mediterranean Sea	DYFAMED	temperate basin	25	2003/03–2007/06	multiple stations	on ship
Scotia Sea	AMLR	polar-continental shelf	171	2004/02–2008/09	multiple stations	on ship
Southern Ocean	WAP	polar-continental shelf	370	2003/01–2011/02	multiple stations	on ship

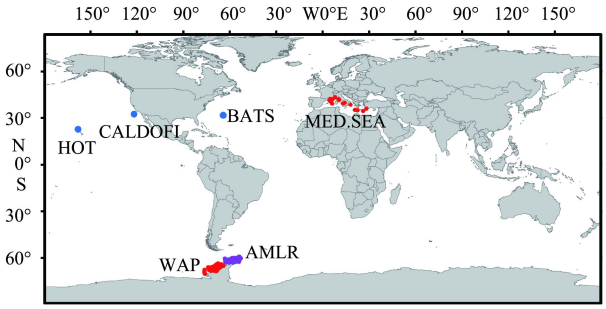

Fig. 1. Locations of the *in situ* NPP measurements: blue dots for the long term monitoring sites, red and purple dots for the ship monitoring sites.

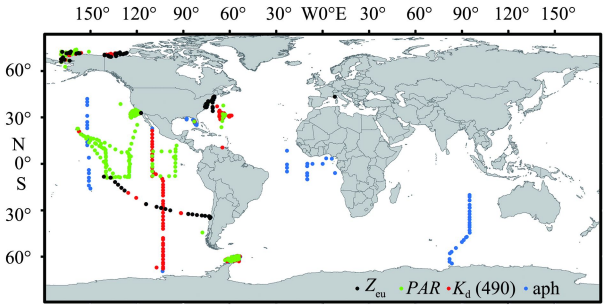

Fig. 2. Locations of the *in situ* data used to assess the input uncertainties of $a_{ph}(\lambda)$ (blue dots), *PAR* (green dots), $K_d(490)$ (red dots) and Z_{eu} (black dots).

Table 2. List of input parameters

Symbol	Definition	Unit
λ	wavelength	nm
z	depth	m
$\bar{a}_{ph}(z)$	phytoplankton pigment absorption coefficient spectrally averaged over 400–700 nm at depth z	m^{-1}
$a_{ph}(0)$	phytoplankton pigment absorption coefficient spectrally averaged at available satellite bands, just below sea surface at depth 0	m^{-1}
Φ	quantum yield of phytoplankton photosynthesis	mol/mol (carbon: photons)
Φ_{max}	maximum value for quantum yield	mol/mol (carbon: photons)
K_m	irradiance where Φ reaches the half of its maximum value	μW/(cm²·nm)
$E_d(\lambda, z)$	downwelling solar irradiance at depth z and wavelength λ	μW/(cm²·nm)
$E(z)$	integration of downwelling solar irradiance $E_d(\lambda, z)$ over wavelength range 400–700 nm at depth z	μW/cm²
v	photoinhibition parameter	$(\mu W/cm^2)^{-1}$

$$E(z) = \int_{400}^{700} E_d(\lambda, 0) \times \exp(-K_d(\lambda) \times z)\, d\lambda, \quad (2)$$

where $K_d(\lambda)$ is the spectral diffuse irradiance attenuation coefficient, which can be calculated from $K_d(490)$ using an empirical model (Austin and Petzold, 1986):

$$K_d(\lambda) = M(\lambda)(K_d(490) - K_w(490)) + K_w(\lambda), \quad (3)$$

where $M(\lambda)$ is a constant and $K_w(\lambda)$ is the spectral diffuse irradiance attenuation coefficient of pure water. In Eq. (2), $E_d(\lambda, 0)$ is the solar irradiance at the sea surface, which can be obtained from the cloud-corrected surface *PAR* via the band decomposition method (Ricchiazzi et al., 1998):

$$E_d(\lambda, 0) = PAR(\lambda, 0) = A \times PAR(0), \quad (4)$$

where $PAR(0)$ is the sea surface *PAR* obtained from the MODIS data, and A is a constant.

The phytoplankton pigment absorption coefficient spectrally averaged over 400 to 700 nm, $\bar{a}_{ph}(z)$, is defined as (Lee et al., 1996):

$$\bar{a}_{ph}(z) = \frac{\int_{400}^{700} a_{ph}(\lambda, z) E_d(\lambda, z)\, d\lambda}{\int_{400}^{700} E_d(\lambda, z)\, d\lambda} \approx \frac{\int_{400}^{700} a_{ph}(\lambda, z)\, d\lambda}{700 - 400}. \quad (5)$$

For satellite data, a rough approximation has been suggested for use (Ma et al., 2014):

$$\bar{a}_{ph}(z) \approx a_{ph}(0), \quad (6)$$

where $a_{ph}(0)$ is the spectrally averaged phytoplankton pigment absorption coefficient just below sea surface at depth 0, which can be calculated from satellite-observed absorption coefficients at available satellite spectral bands:

$$a_{ph}(0) \approx 0.952 \times \frac{\sum_{i=1}^{5} [(a_{ph}(\lambda_{i+1}) + (a_{ph}(\lambda_i)) \times (\lambda_{i+1} - \lambda_i)/2]}{667 - 412}, \quad (7)$$

where i is the MODIS band number and wavelength from λ_1 to λ_5 is 412, 443, 488, 531, 547 and 667 nm, respectively. Here 0.952 was derived from the MODIS remote sensing data in this study by using the same method given by Ma et al. (2014) and Hirawake et al. (2011).

The AbPM model was based on Eqs (1)–(3) and the depth-integrated NPP can be expressed as

$$NPP = \int_0^{Z_{eu}} NPP(z)\, dz, \quad (8)$$

where Z_{eu} is the euphotic zone depth, which can be detected by MODIS. The detailed processing scheme for the input parameters and the whole framework of the model can be found in previous studies (Ma et al., 2014; Lee et al., 1996). Based on Eqs (1)–(4), the AbPM model of NPP can be expressed as

$$NPP = f(a_{ph}(0), K_d(490), Z_{eu}, PAR(0)). \quad (9)$$

The error of AbPM-output of the NPP caused by the four input parameters was quantitatively assessed in this study.

2.4 *Model performance*

The overall performance of AbPM model was evaluated in terms of both bias and variability following the method of Friedrichs et al. (2009). The log normalized bias (B), the root mean square difference (RMSD), and the signed unbiased RMSD (uRMSD) of NPP were calculated as follows:

$$B = \overline{\log_{10}(NPP_{\rm m})} - \overline{\log_{10}(NPP_{\rm d})}, \quad (10)$$

$$RMSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[\log_{10}(NPP_{\rm m}(i)) - \log_{10}(NPP_{\rm d}(i))\right]^2}, \quad (11)$$

$$uRMSD = \mathrm{sign}(\sigma_{\rm m} - \sigma_{\rm d}) \times \sqrt{RMSD^2 - B^2}, \quad (12)$$

where N is the number of samples at each site, $NPP_{\rm m}$ is the modeled NPP, and $NPP_{\rm d}$ represents the *in situ* data at each site. $\sigma_{\rm m}$ and $\sigma_{\rm d}$ are the standard deviations of $\log_{10}(NPP_{\rm m})$ and $\log_{10}(NPP_{\rm d})$, respectively.

2.5 *Uncertainty analysis of input parameters*

For the purpose of this section, we assumed that the AbPM model itself was perfect; therefore, the uncertainties in the model inputs were the only sources of uncertainty in the output. The input uncertainties were assessed using statistical methods which took the simultaneous *in situ* data as the reference values. For each input parameter, the discrepancy (δ) between a model output and the coincident reference value was determined. The mean value of discrepancy is represented by the bias (B), and the standard deviations ($\sigma_{\rm m}$ and $\sigma_{\rm d}$) reflect the remaining uncertainty components. Next, we analyzed the characteristics of each discrepancy and tried to find an appropriate frequency distribution as the uncertainty distribution. A detailed explanation of this methodological approach with accompanying notation and equations can be found in the work of Milutinović and Bertino (2011).

Following the method proposed by Milutinović and Bertino (2011), the simple difference (δ), logarithmic difference ($\delta^{\rm LOG}$), bias (B), logarithmic bias ($B^{\rm LOG}$), root mean square difference ($RMSD$) and logarithmic root mean square difference ($RMSD^{\rm LOG}$) of each input parameter were calculated as follows:

$$B = \bar{\delta} = \frac{1}{n}\sum_{i=1}^{n}\delta_i = \frac{1}{n}\sum_{i=1}^{n}(MOD_i - REF_i), \quad (13)$$

$$B^{\rm LOG} = \overline{\delta^{\rm LOG}} = \frac{1}{n}\sum_{i=1}^{n}\left[\log_{10}(MOD_i) - \log_{10}(REF_i)\right], \quad (14)$$

$$RMSD = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\delta_i - B)^2}, \quad (15)$$

$$RMSD^{\rm LOG} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(\delta_i^{\rm LOG} - B^{\rm LOG}\right)^2}, \quad (16)$$

where MOD denotes the satellite observed data for the model parameter and REF denotes the ship measured data of the corresponding parameter.

2.6 *Uncertainty distribution of four input parameters and uncertainty of NPP*

Following the method proposed by Milutinović and Bertino (2011), δ, $\delta^{\rm LOG}$, B, $B^{\rm LOG}$, $RMSD$ and $RMSD^{\rm LOG}$ of each input parameter were calculated. Our screening indicates that the uncertainty distribution of each parameter follows these rules:

$$Z_{\rm eu} \sim N[Z_{\rm eu} - B(Z_{\rm eu}), RMSD(Z_{\rm eu})], \quad (17)$$

$$PAR \sim N[PAR - B(PAR), RMSD(PAR)], \quad (18)$$

$$\log_{10}(a_{\rm ph}(0)) \sim Nor\left[\log_{10}(a_{\rm ph}(0)) - B^{\rm LOG}(a_{\rm ph}(0)), RMSD^{\rm LOG}(a_{\rm ph}(0))\right], \quad (19)$$

$$\log_{10}(K_{\rm d}(490)) \sim Nor\left[\log_{10}(K_{\rm d}(490)) - B^{\rm LOG}(K_{\rm d}(490)), RMSD^{\rm LOG}(K_{\rm d}(490))\right], \quad (20)$$

where *Nor* is the symbol for the normal probability distribution; *PAR* is the cloud-corrected surface photosynthetically available radiation. Its mean and standard deviation are given in square brackets. For example, the first item [$\log_{10}(a_{\rm ph}(0))$–$B^{\rm LOG}(a_{\rm ph}(0))$], in right side of Eq. (19), is the mean; the second item [$RMSD^{\rm LOG}(a_{\rm ph}(0)$] is the standard deviation. $\log_{10}(a_{\rm ph}(0))$ denotes the satellite remote sensing data for the phytoplankton pigment absorption coefficient just below sea surface at depth 0 in a grid cell, $B^{\rm LOG}(a_{\rm ph}(0))$ denotes the mean value of logarithmic difference ($\delta^{\rm LOG}$), between the satellite observations and the ship measured *in situ* data defined by Eq. (14). The standard deviation for each grid cell is approximately replaced by the [$RMSD^{\rm LOG}(a_{\rm ph}(0))$] obtained from Eq. (16).

The Monte Carlo simulation is the most widely used approach to evaluate the uncertainties propagated through a model (Milutinović and Bertino, 2011). The core of the method is the random sampling from pre-established uncertainty distributions of the input parameters. The model was calculated repeatedly using the random input parameters to obtain an uncertainty distribution for the model output. The uncertainty of model output was related to the number of the random input data (Milutinović and Bertino, 2011).

In this paper, the uncertainty of input parameters due to randomicity was simulated based on the normal probability distribution (i.e., the pre-established uncertainty distribution), denoted by Eq. (17) through Eq. (20). For each bin (i.e., grid cell), 1 200 random data points were simulated as pre-established uncertainty distribution of the input parameters by using computer software. During the process of input parameter simulation, some generated unreliable values (i.e., $Z_{\rm eu}$<5 m or >180 m) were removed (Morel and Maritorena, 2001). Because the input parameters could not be negative, simulated negative values were also removed. After this quality control, a bin of each individual dataset that contained less than 1 140 simulated data points (i.e., less than 95%, compared to 1 200 data points) was also abandoned. A total of 2.1% of the simulated datasets were disregarded in this manner.

Based on the Eq. (13) through Eq. (20), the uncertainty of NPP obtained via AbPM model can be expressed as

$$\Delta NPP_i = NPP - \overline{MNPP} = NPP - \frac{\sum_{j=1}^{n} MNPP_{ij}}{n}, \quad i = 1,2,3,4,5, \quad n = 1\,200, \quad (21)$$

where $MNPP$ is calculated from AbPM using the Monte Carlo simulated random arrays and NPP is retrieved from AbPM using remote sensing data. $\overline{MNPP}$ denotes the mean value for uncertainty distributions of $MNPP$, where $\overline{MNPP}$ represents the NPP value without bias in the input quantities, because all four quantities had been corrected for bias before the Monte Carlo simula-

tions were carried out. In other words, *NPP* incorporates the effects of bias in input quantities, whereas $\overline{MNPP}$ does not. Therefore, the difference (ΔNPP) between $\overline{MNPP}$ and *NPP* is defined as the error caused by input parameter uncertainty. Here, i=1, 2, 3, 4 represents one of the four parameters of Eq. (17) through Eq. (20) involved in one simulation, ΔNPP_i (i=1, 2, 3, 4) reflects the uncertainties brought by single parameter to the AbPM model involved in one simulation, ΔNPP_5 reflects the uncertainties brought by all the parameters to the AbPM model. Therefore, ΔNPP indicates how greatly the NPP is over- or under-estimated.

The percentage bias is defined as the ratio of ΔNPP_i (i=1–5) to the NPP retrieved by remote sensing data. The magnitude of percentage bias reflects the magnitude of the uncertainties brought by each or all the parameters to the AbPM model compared with the NPP retrieved via remote sensing. The positive or negative percentage bias denotes whether this uncertainty increases or decreases the NPP. The coefficient of variation (CV) is defined as the ratio of the standard deviation to the mean. The magnitude of CV reflects the width of the scatter range of the uncertainty brought by the parameters to AbPM model, compared to the mean. Both the percentage bias and CV indicate the uncertainty in the NPP brought by each or all parameters to AbPM model.

2.7 *Correction related to SST*

According to Eppley (1972) and Kishi et al. (2007), the growth rate of phytoplankton is related to temperature as follows (Kishi et al., 2007; Eppley, 1972):

$$\xi(T) = a \times \exp(kT), \tag{22}$$

where T is temperature (°C) and a and k are the coefficients derived from the function of Eppley (1972). In this study, this function was used when the temperature was lower than 10°C, mainly in the low temperature area in the Southern Ocean region. The revised NPP model can be expressed as

$$NPP_{\mathrm{SST}} = NPP \times \xi(T), \quad \text{when } T < 10^{\circ}\mathrm{C}. \tag{23}$$

The calculation result of the revised model was only used to find out the influence of temperature on the AbPM model. And the result was not used for validation and uncertainty analysis of AbPM model.

3 Experimental

3.1 *Validation*

The NPP calculated via AbPM was compared with those calculated by the other two existing models, Vertically Generalized Production Model (VGPM) and Carbon-based Production Model (CbPM) (Westberry et al., 2008; Behrenfeld and Falkowski, 1997). The VGPM model and CBPM model are the commonly used model for monitoring the primary productivity. Related products can be downloaded from the following Web site: http://www.science.oregonstate.edu/ocean.productivity/. The accuracy of the three models was verified by comparing them with the *in situ* data at six sites—HOT, BATS, CALCOFI, DYFAMED, AMLR and WAP (Table 1).

A histogram of the statistical parameters used for comparison among the three models, AbPM, CbPM and VGPM is shown in Fig. 3, which suggests that the overall performance of the AbPM was the best at the long-term monitoring sites. As shown in the figure, AbPM greatly outperformed the other two models at three long-term monitoring sites (HOT, BATS, and CALCOFI) and had the best accuracy, including the smallest values for RMSD, uRMSD and Bias for the NPP estimate, which were obtained from Eqs (10)–(12), and the best correlation between the observations and the modeled productivity.

It is noteworthy that all three sites are located in the Northern Hemisphere. However, at another three ship monitoring sites (DYFAMED, AMLR and WAP), the three models performed similarly, as shown in Fig. 3. Among the three sites, AMLR and WAP sites are located in the Southern Ocean. The accuracy of the

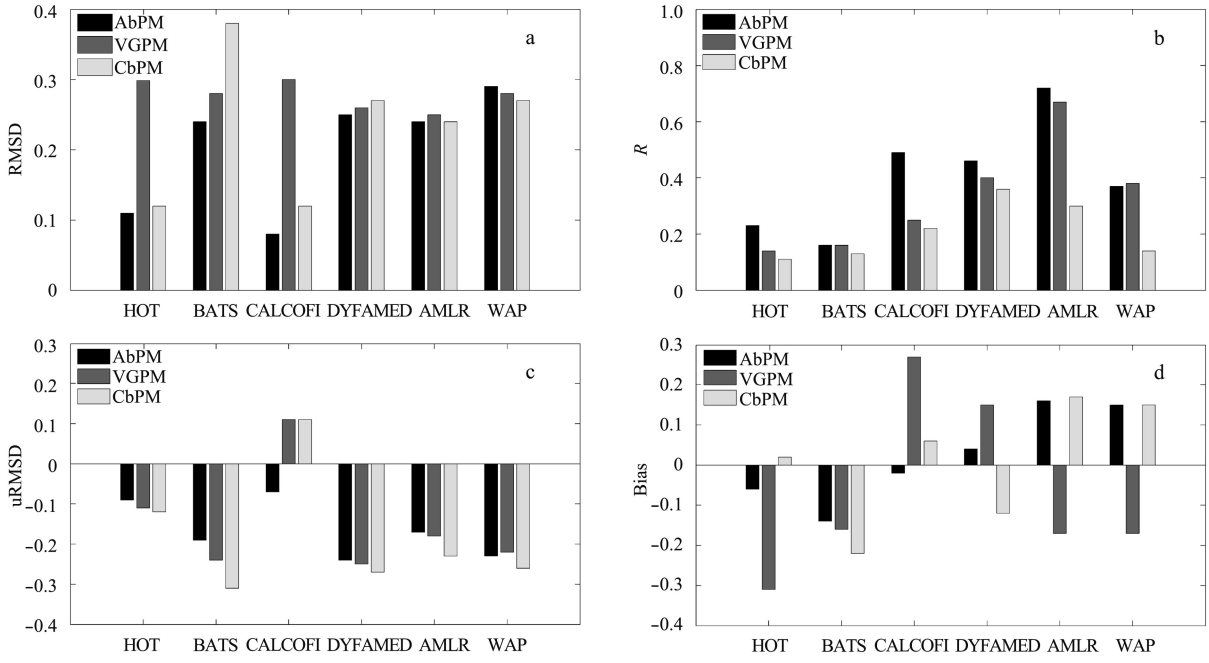

Fig. 3. Histogram of statistical parameters RMSD (a), R (b), uRMSD (c), and Bias (d) used for comparison among the three models: AbPM (black), CbPM (gray) and VGPM (white), respectively. Here RMSD, uRMSD and Bias were obtained from Eqs (10)–(12), and R denotes the correlation between the modeled NPP and observed NPP.

model is obviously lower, especially at the AMLR and WAP sites compared to those at the three sites that had long-term monitoring. The maximum RMSD for the three models at three ship monitoring sites with long-term monitoring was approximately 0.29, however, the RMSD for AbPM averaged at three long-term monitoring sites was lower at 0.18.

The monthly averaged NPPs for the three models in January (summer in the Southern Hemisphere) near the Antarctic Circle are shown in Fig. 4, which suggests that the NPP derived from the AbPM model near the Antarctic Circle was obviously higher than those derived from the other two models. This finding also shows that some uncertainty factors may lead to instability of the model.

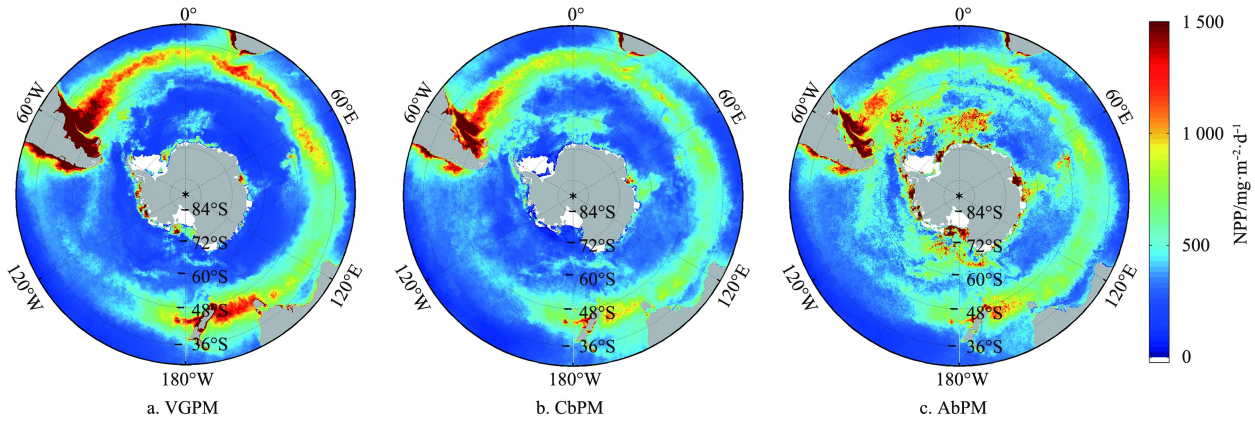

Fig. 4. Depth-integrated NPP (mg/(m^2·d), calculated by carbon) from three models (AbPM, CbPM and VGPM) for the Southern Ocean in January. Values are climatological means for the periods from 2003 to 2012.

3.2 *Uncertainty analysis*

A Monte Carlo simulation was used to model the propagation of uncertainties. First, the accuracy of the input parameters was assessed by using *in situ* data. The bias (B) or logarithmic bias (B^{LOG}) and RMSD or logarithmic RMSD ($RMSD^{\mathrm{LOG}}$) for each parameter are listed in Table 3 and shown in Figs 5a, c, e, and g. The bias (B) or logarithmic bias (B^{LOG}), and the root mean square difference (RMSD) or logarithmic root mean square difference ($RMSD^{\mathrm{LOG}}$) for the input parameters calculated by the method of Milutinović and Bertino (2011) are shown at the top-left. The frequency histograms of simple difference (δ), and logarithmic difference (δ^{LOG}) between model estimates and matching reference *in situ* data confirm that the two input parameters (Z_{eu} and PAR) approximate the Normal Distribution, as shown in Figs 5d and h, and another two input parameters (a_{ph} and K_{d} (490)) approximate the logarithmic Normal Distribution, as shown in Figs 5b and f. In these figures, the frequency distribution of each variable is shown as a frequency histogram of the simple difference (δ), or the logarithmic difference (δ^{LOG}).

Table 3. The uncertainties of input parameters that will be used for assessing the uncertainties propagated through AbPM

Input parameter	Scale	Count	Bias	RMSD
a_{ph}	logarithmic	319	0.071	0.218
Z_{eu}	linear	155	1.927	8.014
K_{d}(490)	logarithmic	444	0.002 4	0.174
PAR	linear	342	1.319	4.240

Based on above experimental method and the global remote sensing data averaged during 2003–2012, the probability distribution for each parameter was obtained by using Eqs (17)–(20). The satellite remote sensed NPP, as denoted by Eq. (8) and the difference ΔNPP_5, as denoted by Eq. (21), were calculated using a random dataset and a Monte Carlo simulation. Figures 6a and b show maps of NPP for January and July, illustrating typical findings for boreal winter/austral summer and boreal summer/austral winter, respectively. Because ΔNPP_5 reflects the uncertainties brought by all parameters to AbPM model, from Figs 6c and d we can deduce that most of NPPs were overestimated by the Monte Carlo simulation and underestimates occurred only sporadically.

Table 4 contains the percentage bias (PB) and coefficient of variation (CV) of the NPP uncertainty brought by each or all the parameters to AbPM model for the global ocean and the coastal zone of Antarctica, respectively. As shown in Table 4, the percentage bias (PB) of the NPP uncertainty brought by all parameters had annual mean of 5.5% for the global ocean. Calculation indicated that, for the global ocean, over 90% of the percentage bias (PB) of the NPP uncertainty brought by all the parameters covered a range of –2%–9% and the total of the percentage bias (PB) of the NPP uncertainty brought by all the parameters covered a range of –5%–15%. As shown in Table 4, the CV brought by all the parameters had annual mean of 105% for the global ocean. Calculation indicated that, for global ocean, the total values of CV brought by all the parameters covered a range of 98%–134%. We also calculated the PB and CV values for the coastal zone of Antarctica. The annual mean of the percentage bias (PB) brought by all the parameters in coastal zone of Antarctica was 7.1%, and this percentage bias (PB) reached the highest value over 7.5% in summer of the southern hemisphere. The CV brought by all the parameters in the coastal zone of Antarctica also increased slightly compared to that in the global ocean as mentioned above. The CV values of coastal zone of Antarctica had an annual mean of 121% covered a range of 118% to 134%.

After assessing the total contributions of all of the parameters (Z_{eu}, a_{ph} (0), K_{d} (490) and PAR), the individual contribution of each input parameter to the uncertainty was also estimated. As mentioned above, a Monte Carlo simulation was used to generate random data for an input parameter following a predetermined frequency distribution (Fig. 5), and the other three input parameters were kept fixed for each repeated calculation. In this way, partial uncertainties for each input parameter were obtained. The percentage bias and CV caused by each individual parameter for the global ocean and the coastal zone of Antarctica were calculated and are listed in Table 4.

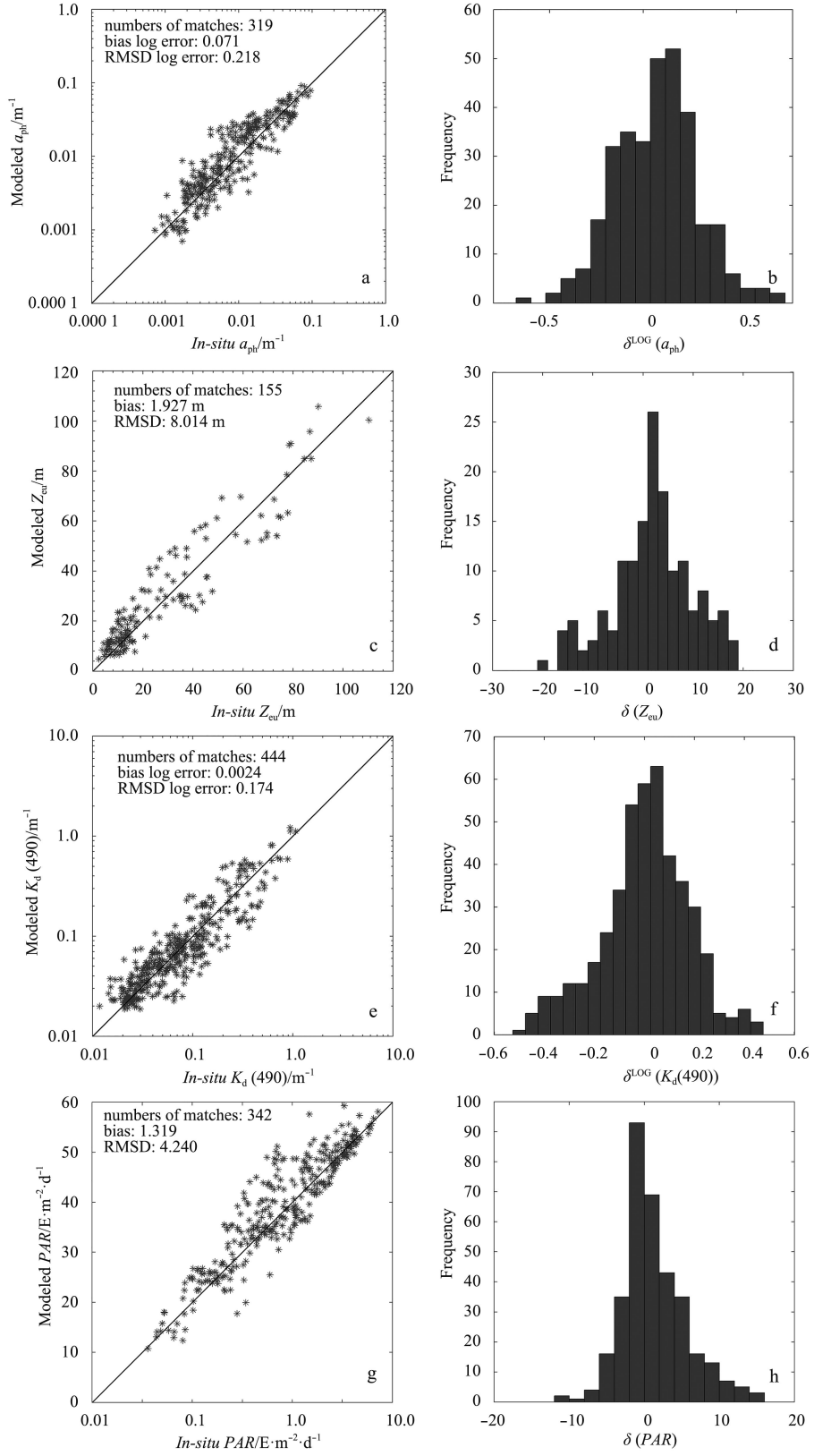

Fig. 5. Comparison of input parameter obtained from satellite observations with reference *in situ* data: a_{ph} (a), Z_{eu} (c), K_d (490) (e) and *PAR* (g). The bias (B) or logarithmic bias (B^{LOG}), and root mean square difference (RMSD) or logarithmic root mean square difference ($RMSD^{LOG}$) of input parameters are shown at the top-left. The frequency histogram of simple difference (δ), and logarithmic difference (δ^{LOG}) between satellite observations and matching reference *in situ* data: a_{ph} (b), Z_{eu} (d), K_d (490) (f) and *PAR* (h), which are implying the normal distribution of two input parameters, Z_{eu} and *PAR*, and the logarithmic normal distribution of another two input parameters, a_{ph} and K_d (490) in AbPM.

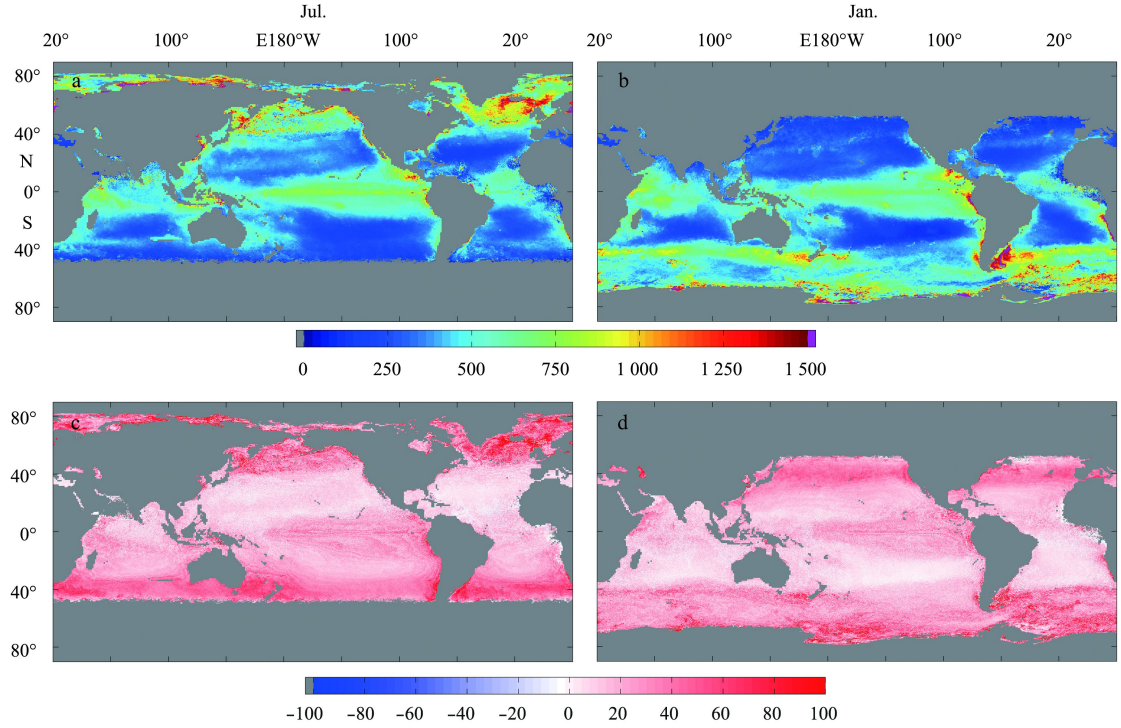

Fig. 6. Satellite remote sensed NPP (mg/(m²·d), calculated by carbon) estimated by the AbPM using MODIS data averaged during 2003–2012 in July (a) and January (b), and the difference ΔNPP_5 (mg/(m²·d), calculated by carbon) between MNPP and NPP are obtained from Eq. (21) in July (c) and January (d). Here ΔNPP_5 denotes the error caused by uncertainty of all the four input parameters. Gray color represents locations with no available remote sensing observations, continent, continental shelf and bins with unreliable statistics.

Table 4. The percentage bias (PB) and coefficient of variation (CV) of the NPP uncertainty brought by each or all the parameters to AbPM model for the global ocean (GO) and the coastal zone of Antarctica (CA), respectively. The uncertainty related to magnitude is denoted by the percentage bias and the uncertainty related to scatter range is denoted by CV

Input parameter	Component of NPP uncertainty			
	Total PB (GO)	Total CV (GO)	Total PB (CA)	Total CV (CA)
Total	5.5%	105%	7.1%	121%
Input parameter	Partial PB (GO)	Partial CV (GO)	Partial PB (CA)	Partial CV (CA)
a_{ph}	4.1%	68%	5.5%	86%
Z_{eu}	1.1%	90%	1.2%	85%
$K_d(490)$	–2.1%	73%	–1.6%	79%
PAR	2.6%	56%	2.4%	54%

3.3 *SST Correction*

With the method in Section 2.7, SST was introduced into the calculation for the coastal zone of Antarctica to revise the influence of low temperature on the AbPM model. Daily MODIS SST products were used in the calculation of AbPM model. Then, the bias and the RMSD of modified NPP were obtained to verify the accuracy. According to our calculation, the RMSD of the SST revised NPP at WAP and AMLR was less than 0.24 while that from unrevised AbPM model was over 0.29. At the same time, the bias from revised model also decreased from to 0.18 to 0.13.

4 Discussion

A comprehensive introduction to the AbPM model was previously given and the accuracy of the AbPM was further verified by *in situ* data for many different ocean basins. As shown in Fig. 3, the bias of NPP estimated by AbPM model is less than those by the existing models (CbPM, VGPM), and the RMSD is also relatively lower. Compared with the two other models, the AbPM has an obviously better NPP estimate. Nevertheless, the AbPM model also has issues in some regions such as in the coastal zone of Antarctica. During the summer in the southern hemisphere, the NPP values in the coastal zone of Antarctica estimated by the AbPM model were higher than those estimated by the other two models (Fig. 4).

To gain more reliable information regarding the applicability of AbPM model, further analysis of AbPM model was performed. A survey on the uncertainty of the four input parameters was completed based on worldwide field data. The input parameter uncertainties were assessed using statistical methods and their bias and RMSD values are listed in Table 3 and shown in Fig. 5. Figure 5 also shows a comparison of four input parameter values obtained from satellite observations with reference *in situ* data, and a frequency histogram of the simple difference (δ), and the logarithmic difference (δ^{LOG}) between the satellite observations and matching reference *in situ* data, which approximates the Normal Distribution. These frequency histograms confirm that

a_{ph} and K_d (490) are very close to the logarithmic normal distribution, and Z_{eu} and PAR are very close to the normal distribution. The distribution histograms for each input parameter are shown in Fig. 5.

Following the predetermined probability density functions, the quantitative uncertainty of the AbPM model output propagated from the input parameters was calculated using the Monte Carlo simulation. Table 4 shows that the percentage bias (PB) had an annual mean of 5.5% for the global ocean which denoted that the output value of AbPM was slightly overestimated. And the coefficient of variation (CV) of the model was 105%, which denoted the scatter range of uncertainty of NPP estimates. These values apply to a period of one year and 9 km×9 km bin size. Present NPP algorithms have met challenge when they were used in the coastal waters (Saba et al., 2011). In this paper, we selected the coastal zone of Antarctica for experiment. The statistical approach for coastal zone of Antarctica applied here relies on the reference data (WAP and AMLR). For the coastal zone of Antarctica, the annual mean of percentage bias (PB) was increased to 7.1%, the annual mean of CV was also increased to 121% significantly. According to the above analysis, the adverse effect of bias in NPP estimates derived from remote sensing data is obvious. Therefore, it is importance to identify and neutralize the sources of such a bias.

Next, the uncertainties of AbPM model from the four input parameters were separately assessed. According to our calculation (Table 4), the partial percentage bias of AbPM caused by a_{ph} (0), Z_{eu}, K_d (490) and PAR were 4.1%, 1.1%, –2.1% and 2.6%, respectively. Compared with the other three parameters, a_{ph} plays more important role in the uncertainties of AbPM model. In addition, partial CV resulting from the uncertainty of Z_{eu} was the highest. Then, we have assessed the uncertainties caused by the four input parameters in the coastal zone of Antarctica. Table 4 lists the partial percentage bias (PB) and partial CV originating from uncertainties in the input parameters. It can be seen that, partial percentage bias (PB) and partial CV have no obvious changes for Z_{eu}, K_d (490) and PAR. At the same time, the partial percentage bias of AbPM caused by a_{ph} rose from 4.1% to 5.5%, and the corresponding CV values also rose from 68% to 86%. These changes indicated that the increase in the total bias for the coastal zone of Antarctica was primarily caused by a_{ph}, and it also made a major contribution to the increase of total CV.

The above analysis indicates that the accuracy of the a_{ph} principally affects the bias of the NPP model. Several potential problems may cause a lower precision of a_{ph}. In Eq. (5), the $\bar{a}_{ph}(z)$ is considered to be approximately equal to a_{ph} (0), a phytoplankton pigment absorption coefficient just below sea surface at depth 0, with ignoring vertical variations. This approximation may result in errors of the estimated phytoplankton pigment absorption at different depths (Gordon and Clark, 1980; Lee and Carder, 2004; Ma et al., 2014). On the other hand, the high CV indicates that the retrieved NPP was sensitive to the scatter range of the distribution of input parameters. Because the field data of input parameters were selected from a wide range, the corresponding standard deviation was relatively higher. This property may cause a higher CV. Therefore, a more reasonable and accurate a_{ph} data should be valuable and corresponding processing method should be developed. Although the other three input parameters have less effect on the model than a_{ph}, higher quality data are still needed. At the same time, more field data are necessary to obtain the mean and standard deviation for the four parameters in the different oceans. Thus, the randomicity of input parameters should be known more accurately.

The effect of the input parameters on model has been analyzed, and the results showed that the model's accuracy was affected by the uncertainties of input parameters. But, it did not indicate that the model is only affected by the input parameters. Some researchers found that input uncertainties can explained more than a half of the mismatch between the model estimates and field observations (Friedrichs et al., 2009; Saba et al., 2011; Milutinović and Bertino, 2011). Here, we tried to find out the potential factors that may affect the accuracy of the model. For example, the quantum yield of photosynthesis is an important parameter to AbPM model, but their values are related to the uncertainties of nitrogen concentration. In fact, the quantum yield of photosynthesis can be influenced by iron, phytoplankton size and nutrients (Woźniak et al., 2007; Uitz et al., 2008; Hiscock et al., 2008). Secondly, during the light-saturation season (summer in the coastal zone of Antarctica), the low seawater temperature may also limit phytoplankton photosynthesis (Reay et al., 2001).

To verify our hypothesis, a preliminary experimental study on the effect of SST on NPP retrieval has been completed. According to calculation, he RMSD and bias of NPP at WAP and AMLR were significantly decreased when a correction related to SST was introduced. The investigation confirmed that the SST related correction is effective for improving the model's accuracy in low temperature and light saturated waters.

The uncertainty analysis in Table 4 indicated that the AbPM model has less accuracy in certain oceans such as the coastal zone of Antarctica, compared with in the global ocean. This study was trying to discover the reason. Our study indicated that a_{ph}, as one of four input parameters, played more important role in the uncertainty of the AbPM model output. Furthermore, we also found that some other factors, such as the low temperature, may also be important to the retrieved NPP errors. Future work should be arranged to improve the AbPM model based on the results of this study.

Acknowledgements

We thank the organizations for free data providing. The MODIS ocean color productions and *in situ* data of phytoplankton pigment absorption were downloaded from NASA Ocean Color websites (http://oceancolor.gsfc.nasa.gov/). The *in situ* NPP data were downloaded from the website of Hawaii Ocean Time-series program (http://hahana.soest.hawaii.edu/hot/), Bermuda Atlantic Time-series Study (http://bats.bios.edu/), California Cooperative Oceanic Fisheries Investigations (http://www.calcofi.org/), Biological and Chemical Oceanography Data Management Office (http://www.bco-dmo.org/), Palmer Station Long-Term Ecological Research (http://pal.lternet.edu/data/) and Atmospheric Dynamics and Fluxes in the Mediterranean Sea (http://www.obs-vlfr.fr/cd_rom_dmtt/dyf_main.htm).

References

Antoine D, Morel A. 1996. Oceanic primary production: 1. Adaptation of a spectral light-photosynthesis model in view of application to satellite chlorophyll observations. Global Biogeochemical Cycles, 10(1): 43–55

Austin R W, Petzold T J. 1986. Spectral dependence of the diffuse attenuation coefficient of light in ocean waters. Optical Engineering, 25(3): 253471

Behrenfeld M J, Falkowski P G. 1997. A consumer's guide to phytoplankton primary productivity models. Limnology and Oceanography, 42(7): 1479–1491

Behrenfeld M J, Randerson J T, McClain C R, et al. 2001. Biospheric

primary production during an ENSO transition. Science, 291(5513): 2594–2597

Bianchi A A, Bianucci L, Piola A R, et al. 2005. Vertical stratification and air-sea CO_2 fluxes in the Patagonian shelf. Journal of Geophysical Research, 110(C7), doi: 10.1029/2004JC002488

Carr M E, Friedrichs M A M, Schmeltz M, et al. 2006. A comparison of global estimates of marine primary production from ocean color. Deep Sea Research Part II—Topical Studies in Oceanography, 53(5–7): 741–770

Eppley R W. 1972. Temperature and phytoplankton growth in the sea. Fishery Bulletin, 70: 1063–1086

Falkowski P G, Barber R T, Smetacek V. 1998. Biogeochemical controls and feedbacks on ocean primary production. Science, 281(5374): 200–206

Friedrichs M A M, Carr M E, Barber R T, et al. 2009. Assessing the uncertainties of model estimates of primary productivity in the tropical Pacific Ocean. Journal of Marine Systems, 76(1–2): 113–133

Garcia H E, Locarnini R A, Boyer T P, et al. 2010. World ocean atlas 2009, volume 4: nutrients (phosphate, nitrate, silicate). In: Levitus S, ed. NOAA Atlas NESDIS 71. Washington, DC: US Government Printing Office, 398

Gordon H R, Clark D K. 1980. Remote sensing optical properties of a stratified ocean: an improved interpretation. Applied Optics, 19(20): 3428–3430

Hirawake T, Takao S, Horimoto N, et al. 2011. A phytoplankton absorption-based primary productivity model for remote sensing in the Southern Ocean. Polar Biology, 34(2): 291–302

Hiscock M R, Lance V P, Apprill A M, et al. 2008. Photosynthetic maximum quantum yield increases are an essential component of the Southern Ocean phytoplankton response to iron. Proceedings of the National Academy of Sciences of the United States of America, 105(12): 4775–4780

Kiefer D A, Cullen J J. 1991. Phytoplankton growth and light absorption as regulated by light, temperature, and nutrients. Polar Research, 10(1): 163–172

Kishi M J, Kashiwai M, Ware D M, et al. 2007. NEMURO—a lower trophic level model for the North Pacific marine ecosystem. Ecological Modelling, 202(1–2): 12–25

Lee Z, Lance V P, Shang Shaoling, et al. 2011. An assessment of optical properties and primary production derived from remote sensing in the Southern Ocean (SO GasEx). Journal of Geophysical Research, 116(C4): doi: 10.1029/2010JC006747

Lee Z P, Carder K L. 2004. Absorption spectrum of phytoplankton pigments derived from hyperspectral remote-sensing reflectance. Remote Sensing of Environment, 89(3): 361–368

Lee Z P, Carder K L, Marra J, et al. 1996. Estimating primary production at depth from remote sensing. Applied Optics, 35(3): 463–474

Longhurst A, Sathyendranath S, Platt T, et al. 1995. An estimate of global primary production in the ocean from satellite radiometer data. Journal of Plankton Research, 17(6): 1245–1271

Ma Sheng, Tao Zui, Yang Xiaofeng, et al. 2014. Estimation of marine primary productivity from satellite-derived phytoplankton absorption data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7(7): 3084–3092

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

Marra J, Ho C, Trees C C. 2003. An alternative algorithm for the calculation of primary productivity from remote sensing data. LDEO Technical Report # LDEO-2003–1. Grant: National Aeronautics and Space Administration

Marra J, Trees C C, O'Reilly J E. 2007. Phytoplankton pigment absorption: a strong predictor of primary productivity in the surface ocean. Deep Sea Research Part I—Oceanographic Research Papers, 54(2): 155–163

McClain C R, Cleave M L, Feldman G C, et al. 1998. Science quality Sea WiFS data for global biosphere research. Sea Technology, 39(9): 10–15

Medina-Gómez I, Herrera-Silveira J A. 2006. Primary production dynamics in a pristine groundwater influenced coastal lagoon of the Yucatan Peninsula. Continental Shelf Research, 26(8): 971–986

Milutinović S, Behrenfeld M J, Johannessen J A, et al. 2009. Sensitivity of remote sensing-derived phytoplankton productivity to mixed layer depth: lessons from the carbon-based productivity model. Global Biogeochemical Cycles, 23(4): doi: 10.1029/2008GB003431

Milutinović S, Bertino L. 2011. Assessment and propagation of uncertainties in input terms through an ocean-color-based model of primary productivity. Remote Sensing of Environment, 115(8): 1906–1917

Morel A, Maritorena S. 2001. Bio-optical properties of oceanic waters: a reappraisal. Journal of Geophysical Research, 106(C4): 7163–7180

Muller-Karger F, Varela R, Thunell R, et al. 2001. Annual cycle of primary production in the Cariaco Basin: response to upwelling and implications for vertical export. Journal of Geophysical Research, 106(C3): 4527–4542

Ondrusek M E, Bidigare R R, Waters K, et al. 2001. A predictive model for estimating rates of primary production in the subtropical North Pacific Ocean. Deep Sea Research Part II—Topical Studies in Oceanography, 48(8–9): 1837–1863

Platt T, Sathyendranath S. 1993. Estimators of primary production for interpretation of remotely sensed data on ocean color. Journal of Geophysical Research, 98(C8): 14561–14576

Reay D S, Priddle J, Nedwell D B, et al. 2001. Regulation by low temperature of phytoplankton growth and nutrient uptake in the Southern Ocean. Marine Ecology Progress Series, 219: 51–64

Ricchiazzi P, Yang Shiren, Gautier C, et al. 1998. SBDART: a research and teaching software tool for plane-parallel radiative transfer in the Earth's atmosphere. Bulletin of the American Meteorological Society, 79(10): 2101–2114

Saba V S, Friedrichs M A M, Antoine D, et al. 2011. An evaluation of ocean color model estimates of marine primary productivity in coastal and pelagic regions across the globe. Biogeosciences, 8(2): 489–503

Saba V S, Friedrichs M A M, Carr M E, et al. 2010. Challenges of modeling depth-integrated marine primary productivity over multiple decades: a case study at BATS and HOT. Global Biogeochemical Cycles, 24(3): doi: 10.1029/2009GB003655

Siegel D A, Westberry T K, O'Brien M C, et al. 2001. Bio-optical modeling of primary production on regional scales: the Bermuda BioOptics project. Deep Sea Research Part II—Topical Studies in Oceanography, 48(8–9): 1865–1896

Sunda W G, Huntsman S A. 1997. Interrelated influence of iron, light and cell size on marine phytoplankton growth. Nature, 390(6658): 389–392

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

Westberry T, Behrenfeld M J, Siegel D A, et al. 2008. Carbon-based primary productivity modeling with vertically resolved photoacclimation. Global Biogeochemical Cycles, 22(2): doi: 10.1029/2007GB003078

Woźniak B, Ficek D, Ostrowska M, et al. 2007. Quantum yield of photosynthesis in the Baltic: a new mathematical expression for remote sensing applications. Oceanologia, 49(4): 527–542