

This work was written as part of one of the author's official duties as an Employee of the United States Government and is therefore a work of the United States Government. In accordance with 17 U.S.C. 105, no copyright protection is available for such works under U.S. Law.

Public Domain Mark 1.0

<https://creativecommons.org/publicdomain/mark/1.0/>

Access to this work was provided by the University of Maryland, Baltimore County (UMBC) ScholarWorks@UMBC digital repository on the Maryland Shared Open Access (MD-SOAR) platform.

Please provide feedback

Please support the ScholarWorks@UMBC repository by emailing scholarworks-group@umbc.edu and telling us what having access to this work means to you and why it's important to you. Thank you.

Instantaneous photosynthetically available radiation models for ocean waters using neural networks

KAMAL ARYAL,¹  PENG-WANG ZHAI,^{1,*} MENG GAO,^{2,3} AND BRYAN A. FRANZ³

¹Department of Physics, Joint Center for Earth System Technology, University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, Maryland 21250, USA

²Science Systems and Applications, Inc., Lanham, Maryland 20706, USA

³NASA Goddard Space Flight Center, Code 616, Greenbelt, Maryland 20771, USA

*Corresponding author: pwzhai@umbc.edu

Received 5 September 2022; revised 23 October 2022; accepted 26 October 2022; posted 26 October 2022; published 16 November 2022

Instantaneous photosynthetically available radiation (IPAR) at the ocean surface and its vertical profile below the surface play a critical role in models to calculate net primary productivity of marine phytoplankton. In this work, we report two IPAR prediction models based on the neural network (NN) approach, one for open ocean and the other for coastal waters. These models are trained, validated, and tested using a large volume of synthetic datasets for open ocean and coastal waters simulated by a radiative transfer model. Our NN models are designed to predict IPAR under a large range of atmospheric and oceanic conditions. The NN models can compute the subsurface IPAR profile very accurately up to the euphotic zone depth. The root mean square errors associated with the diffuse attenuation coefficient of IPAR are less than 0.011 m^{-1} and 0.036 m^{-1} for open ocean and coastal waters, respectively. The performance of the NN models is better than presently available semi-analytical models, with significant superiority in coastal waters. © 2022 Optica Publishing Group under the terms of the [Optica Open Access Publishing Agreement](#)

<https://doi.org/10.1364/AO.474914>

1. INTRODUCTION

Instantaneous photosynthetically available radiation (IPAR) is the scalar solar irradiance, E_0 , integrated in the visible electromagnetic spectrum (400–700 nm) [1]. Ocean surface IPAR and its subsurface vertical profile plays a critical role in many biological, chemical, and physical processes in the upper ocean. For instance, radiation in this range is utilized by phytoplankton for photosynthesis in the ocean euphotic zone [2–5] and, hence, constrains the type and distribution of algae species that contribute significantly to primary production [6]. It is critical to understand the physiological response of phytoplankton to light availability, such as change in cellular pigmentation [7], solar-stimulated fluorescence [8], fluorescence quantum yield [9,10], and photochemical and non-photochemical quenching processes [11],

$$\text{IPAR} = \frac{1}{hc} \int_{400}^{700} \lambda E_0(\lambda) d\lambda. \quad (1)$$

Photosynthesis by phytoplankton contributes to nearly half of the biosphere's net primary production [12]. The main variables needed in satellite-based primary production models [13,14] are chlorophyll-a concentration ([Chl_a]), vertical distribution of photosynthetically available radiation up to

euphotic zone depth [15], and sea surface temperature. IPAR also determines heat budget in the upper ocean [16–19] and water visibility [20,21]. Thus, accurate estimation of IPAR at the surface and its vertical distribution in different oceanic and atmospheric conditions is extremely important for the estimation of net primary production on the global scale.

Surface IPAR depends mainly on the solar zenith angle (SZA), aerosol/cloud loading, ozone concentration, and ocean surface wind speed, while attenuation below the surface depends on inherent optical properties (IOPs) of ocean waters. In clear open ocean waters, the IOPs can be parameterized by chlorophyll-a concentration. In more turbid coastal waters, absorption and scattering by suspended sediments and color dissolved organic matter (CDOM) should also be considered [22,23]. Surface IPAR can be measured *in situ* or modeled by using numerical algorithms such as those from [24,25], SBDART [26], and [27]. These models are based on radiative transfer models with the plane-parallel approximation, whose performance rapidly decreases at high SZAs with relative uncertainty up to 55% [28]. This makes it necessary to consider the spherical shell effect when calculating IPAR in polar regions with low solar elevation.

The attenuation of IPAR below the ocean surface can be modeled by the diffuse attenuation coefficient (K_d). Empirical and semi-analytical approaches have been used to calculate K_d for different water types. Empirical algorithms of K_d are based on the regression analysis of the diffuse attenuation coefficient of PAR with [Chla] concentration [29–32], which are strongly dependent on the field datasets used in the model development and are only applicable for open ocean waters. It should also be noted that these empirical algorithms [31,32] are only suitable for $K_d(490)$ values up to 0.3 m^{-1} [33]. Other semi-analytical models [34–36] use the absorption coefficient and backscattering coefficient as inputs. These semi-analytical models work better than the empirical models for open ocean waters and mildly turbid coastal waters, though they have been found to underpredict the diffuse attenuation coefficient in highly turbid coastal waters. The authors in [33] have shown that the semi-analytical model [34] underestimates $K_d(490)$ by a factor of 2–3 in comparison with *in situ* measurements in the Chesapeake Bay region.

In addition, these semi-analytical models [34,35] do not consider the effect of backscattering fraction in K_d . Using *in situ* measurements in Upper Chesapeake Bay, the authors of [37] have shown that K_d from the surface to certain depths in highly turbid coastal waters has significant dependence on backscattering fraction. These limitations indicate a need for accurate and efficient models to predict underwater light fields in turbid coastal waters. Moreover, currently there is no standalone model that can compute surface IPAR and its vertical profile below the surface considering both atmospheric and oceanic conditions as inputs.

In this work, we present two IPAR prediction models, based on neural networks (NNs), which can be used to compute accurate surface IPAR and its vertical profile up to euphotic zone depth, considering both atmospheric and oceanic conditions as inputs. Variation of backscattering fractions and the spherical shell effect [38] are considered while preparing the training datasets. One model is for open ocean and the other for coastal waters. The IPAR profile below the surface is modeled using an exponential decay function with two free parameters as in [35]. Both NN models are trained with a large volume of data spanning a wide range of atmospheric and ocean conditions, which makes them suitable for most applications. The main advantages of these NN models include their accuracy, robustness to noise, and computational efficiency. The estimated IPAR values can be integrated to find daily PAR, and the diffuse attenuation coefficient of IPAR (K_{IPAR}) can be converted to the diffuse attenuation coefficient of daily PAR (K_{PAR}) using the technique developed by [39].

Our NN models perform better than the semi-analytical models in both open ocean waters and coastal waters; thus, they can be used as a replacement to overcome the limitations of semi-analytical models. This paper is organized as follows: Section 2 defines the data and methodology used in this work, Section 3 presents the results and discussion, and Section 4 summarizes the conclusion reached from this study.

2. MATERIALS AND METHODS

A. Data

We simulated 8000 cases each for open ocean and coastal waters models using a vector radiative transfer model based on the successive order of scattering method (RTSOS) [40,41]. RTSOS is a plane-parallel radiative transfer model that can simulate a polarized radiation field in the combined atmosphere and ocean system. RTSOS uses the improved pseudo-spherical shell (IPSS) algorithm to treat the spherical shell effects [38], which makes our dataset applicable to high latitude regions. To calculate IPAR, RTSOS is used to loop over the visible spectrum from 400 nm to 700 nm and integrate the scalar irradiance. For details of the calculation, readers are referred to [11]. The simulated cases represent different atmosphere and ocean conditions with a large range of parameters. The atmospheric and oceanic parameters are uniformly sampled within prescribed bounds using the Halton quasi-random number sequence [42,43], which is of low discrepancy and has more uniform coverage over the domain of parameter ranges in multidimensional simulation compared to other pseudorandom uniform numbers [44].

The bio-optical model of open ocean waters is represented by chlorophyll-a concentration uniformly distributed from 0.001 to 30 mgm^{-3} in logarithmic scale. For a given [Chla] value, the ocean IOPs are determined in the same way as [45]. Specifically, the ocean water absorption coefficient includes contributions from pure sea water [46–49], [Chla] covariant particles [50], and CDOM [30]. The scattering coefficient and backscattering fraction of particulate matter is parameterized by [Chla] [30]. The phase function of the particulate matter is determined by the backscattering fraction [51,52]. The total phase function is calculated by mixing the phase functions of the pure sea water and particulate matters weighted by their scattering coefficients. The reduced scattering matrix of ocean waters is based on [53].

The ocean surface roughness is parameterized in terms of wind speed [54], which ranges from 0.5 to 10 ms^{-1} . The SZA is between 0 and 85 degrees. The aerosol model is randomly chosen among [55] and the ocean color aerosol model of [56]. The aerosol optical depth (AOD) at 550 nm ranges within 0 – 0.5 , relative humidity (RH) varies from 50 – 95% . The optical properties of aerosols are mixed with the Rayleigh atmosphere to get the total values. Gas absorption due to H_2O , CO_2 , O_2 , CH_4 , ozone, and NO_2 is considered (see [45] for more details), though only ozone amount is allowed to vary in the synthetic dataset. The range of parameters for open ocean waters are summarized in Table 1.

Coastal waters are represented by the seven-parameter bio-optical model as described in [57,58]. In this bio-optical model, coastal waters are assumed to consist of four components: pure sea water, phytoplankton particles, non-algae particles (NAPs), and CDOM. The parameterization of the IOPs are [57]

$$a_{\text{ph}}(\lambda) = A_{\text{ph}}(\lambda)[\text{Chla}]^{E_{\text{ph}}(\lambda)}, \quad (2)$$

$$a_{\text{dg}}(\lambda) = a_{\text{dg}}(440) \exp[-S_{\text{dg}}(\lambda - 440)], \quad (3)$$

$$b_{\text{bp}}(\lambda) = b_{\text{bp}}(660) [\lambda/660]^{-S_{\text{bp}}}, \quad (4)$$

Table 1. Range of Parameters for Radiative Transfer Simulations for Open Ocean Waters

Parameters	[Chla]	Wind Speed	Ozone	SZA	AOD	Aerosol Models	RH
Unit	mgm ⁻³	ms ⁻¹	Dobson unit	Degrees			%
Range	0.001–30	0.5–10	150–450	0–85	0–0.5	Rural, Oceanic, Urban, etc.	50–95

$$B_p(\lambda) = B_p(660)[\lambda/660]^{-S_{Bp}}, \quad (5)$$

where wavelength λ is in nm, a_{ph} denotes the absorption coefficient for phytoplankton, a_{dg} is the total absorption coefficient for both NAP and CDOM, and b_{bp} and B_p are the total backscattering coefficient and total backscattering fraction for all particles (phytoplankton and NAP), respectively. Values of the coefficients A_{ph} and E_{ph} are obtained from [50]. S_{dg} is the exponential spectral slope for $a_{dg}(\lambda)$; S_{bp} and S_{Bp} are the polynomial spectral slopes for $b_{bp}(\lambda)$ and $B_p(\lambda)$, respectively. The absorption coefficient $a_w(\lambda)$ and scattering coefficient $b_w(\lambda)$ for pure sea water are from the experimental data [46,59,60]. The backscattering fraction for water is 0.5 [61]. The total absorption coefficient is $a(\lambda) = a_w(\lambda) + a_{ph}(\lambda) + a_{dg}(\lambda)$, and the total backscattering coefficient is $b(\lambda) = b_{bw}(\lambda) + b_{bp}(\lambda)$. The scattering coefficient of the phytoplankton and NAP particles can be derived from $b_p = b_{bp}/B_p$. The scattering coefficient for CDOM is assumed to be negligible. All ocean parameters are assumed to be independent of ocean depth.

Similar to the open ocean water cases, the Fourier–Forand (FF) phase function [51,52] is used to represent the phase function of the ocean water particulate matter, which agrees with various *in situ* volume scattering function (VSF) measurements [62]. The range of oceanic parameters for coastal waters is shown in Table 2. The atmospheric parameters including aerosol models, wind speed, AOD, SZA, and RH are in the same range as in the open ocean case.

B. Methodology

For all open ocean and coastal waters cases, the IPAR profile is simulated from the surface up to 200 m depth with 2 m resolution using RTSOS. To model the subsurface IPAR profile in the NNs, the exponential decay function from [35] is adopted:

$$\text{IPAR}(z) = \text{IPAR}(0^-) e^{-\left(K_1 + \frac{K_2}{(1+z)^{0.5}}\right)z}, \quad (6)$$

where K_1 and K_2 are depth-independent parameters whose values depend on atmospheric and oceanic conditions. For each case, their values are obtained by fitting IPAR values from the surface to euphotic zone depth [3]. From the simulated datasets, we explore the dependence of IPAR just below the surface ($\text{IPAR}(0^-)$), K_1 , and K_2 on atmospheric and oceanic variables using correlation matrix analysis (Figs. 1 and 2). Aerosol properties are characterized by extinction optical depths (τ_{ext}), the

single scattering albedo (ω), and the angstrom exponent (α). τ_{ext} at 550 nm is in the range 0–0.5, and ω at 550 nm is in the range 0.88–0.99. The angstrom exponent is calculated between wavelength 410 nm and 670 nm, yielding a range of –0.06 to 1.45.

Two independent multilayer perceptron (MLP) NN models are designed for open ocean and coastal waters using machine learning PyThon library Keras [63]. An MLP is a feed forward and backpropagated artificial NN [64,65]. A simple form of MLP with one input layer, one output layer, and k hidden layers can be represented as [66,67]

$$h_1 = \phi(W_1^T x + b_1), \quad (7)$$

$$h_{p+1} = \phi(W_{p+1}^T h_p + b_p); \quad p = 1, 2, \dots, k-1, \quad (8)$$

$$y = W_{k+1}^T h_k + b_{k+1}, \quad (9)$$

where x is the input vector containing input parameters. The input parameters for the open ocean NN model include the wind speed, [Chla], SZA (θ_0), extinction AOD (τ_{ext}), single scattering albedo (ω), angstrom exponent (α), and ozone column amount. In addition to these seven parameters, the NN model for coastal waters includes three additional parameters: the total absorption coefficient at 490 nm ($a(490)$), the total backscattering coefficient at 490 nm ($b_b(490)$), and the total backscattering fraction for both phytoplankton and NAP at 670 nm ($B_p(670)$), which are calculated based on the bio-optical model [Eqs. (2)–(5)] for each case. The ranges of these parameters are $a(490) \in (0.015 - 1.274) \text{ m}^{-1}$, $b_b(490) \in (0.001 - 0.116) \text{ m}^{-1}$, and $B_p(670) \in (0.002 - 0.05)$, respectively. y is the three-dimensional output vector of IPAR(0^-), K_1 , and K_2 . W_p and b_p represent the weight and bias matrix for the p^{th} layer. The output from h_p of p^{th} layer acts as an input to $p+1^{\text{th}}$ layer. ϕ denotes the activation function, which operates on the output from each previous layer before feeding it to the next layer. For each training iteration, the NN algorithm first makes a prediction, measures the mean squared error, and then goes back through each layer in reverse to minimize errors by optimizing weights and bias matrices using the Adam optimization algorithm.

We randomly selected 90% of the dataset from each water type for training and validation, with the remaining 10% withheld for testing purposes. Of the training dataset, 80% of the

Table 2. Range of Parameters for Radiative Transfer Simulations for Coastal Waters

Parameters	[Chla]	$a_{dg}(440)$	$b_{bp}(660)$	$B_p(660)$	S_{dg}	S_{bp}	S_{Bp}
Unit	mgm ⁻³	m ⁻¹	m ⁻¹		nm ⁻¹	nm ⁻¹	nm ⁻¹
Range	0.001–30	0–2.5	0–0.1	0–0.05	0.01–0.02	0–0.5	–0.2 – 0.2

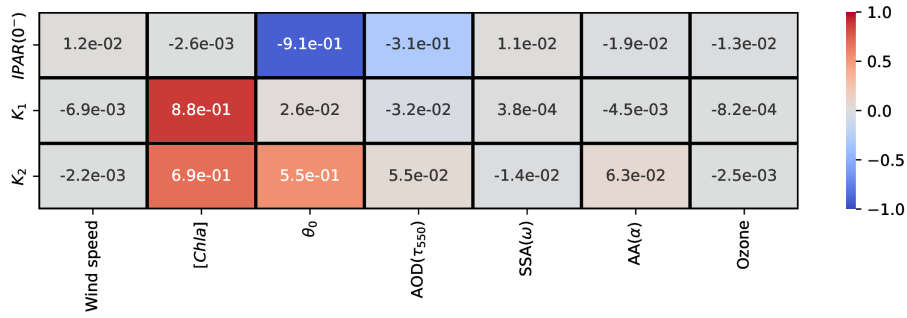


Fig. 1. Correlation of IPAR(0^-), K_1 , and K_2 with different parameters for open ocean waters.

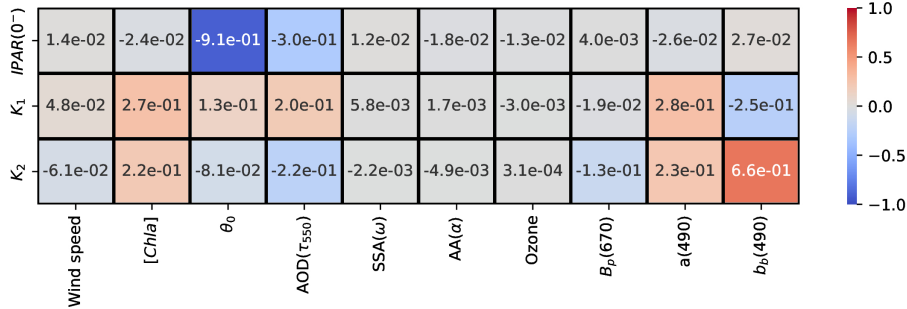


Fig. 2. Correlation of IPAR(0^-), K_1 , and K_2 with different parameters for coastal waters.

dataset was used for training, and 20% was used for validation. After testing several NN architectures with different activation functions [rectified linear unit (ReLU) and Leaky-ReLU, sigmoid, and hyperbolic tangent] and learning rates, we selected a model having two hidden layers with 128 and 64 neurons, respectively, for open ocean waters and a model having three hidden layers with 100, 64, and 64 neurons, respectively, for coastal waters as the best performing models. In hidden layers of NN models, the ReLU activation function is found to work best for open ocean, and hyperbolic tangent (Tanh) is found to work best for coastal waters, which are defined as

$$\text{ReLU}(z) = \max(0, z). \quad (10)$$

$$\text{Tanh}(z) = \frac{(e^z - e^{-z})}{(e^z + e^{-z})}, \quad (11)$$

The linear activation function is used in the output layer of both models. To efficiently train the NN, we performed the following operations, which further optimized the training strategy discussed by [67]:

- Input data are scaled in the range of [0,1] using the minimum and maximum values of each input variable from the training dataset.
- Output data are normalized by dividing their standard deviation of the training dataset.
- The exponential decay learning rate schedule is used. The learning rate is reduced by a factor of 10 every 150 epochs. The initial learning rate of 0.01 is used in open ocean model, and 0.005 is used in coastal waters model.
- The batch size of 100 is used.

- The early stopping technique is used to decrease model training time and prevent overfitting. The model training was stopped when the model performance did not improve for 100 epochs, and the best weights and bias are restored.

The NN models predict IPAR(0^-), K_1 , and K_2 , which can be used in Eq. (6) to represent the IPAR vertical profile up to euphotic zone depths. The IPAR vertical profile can be used to calculate the diffuse attenuation coefficient at any depth using the following equation:

$$K_{NN}(z) = \frac{-1}{z} \ln \left[\frac{\text{IPAR}(z)}{\text{IPAR}(0^-)} \right], \quad (12)$$

which is introduced here in order to compare with the attenuation coefficient $K_{\text{Lee}}(z)$ from semi-analytical model (hereafter referred to as Lee model) developed by [35].

C. Statistical Metrics

To evaluate the model performance, we used the test dataset and three statistical error matrices: coefficient of determination, mean absolute error (MAE), and root mean square error (RMSE), which are defined below.

Coefficient of determination (R^2) is equal to the square of the Pearson correlation coefficient between the predicted and true values, and it is defined as [68]

$$R^2 = \left[\frac{n \sum xy - (\sum x)(\sum y)}{\sqrt{n \sum x^2 - (\sum x)^2} \sqrt{n \sum y^2 - (\sum y)^2}} \right]^2. \quad (13)$$

It describes the proportion of the variance in the dependent variable that can be explained by variance in the independent variable. A high R^2 value indicates good model performance.

MAE and RMSE are defined as

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |P_i - T_i|, \quad (14)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n [P_i - T_i]^2}, \quad (15)$$

where P_i is the value predicted by the NN model and T_i is the true value from the RTSOS simulation. RMSE and MAE measure the residual errors in the output units. A linear regression line is also fitted between the predicted and true values in the form of $y = mx + c$, where m is the slope and c is the y intercept.

3. RESULTS AND DISCUSSION

A. Results

To determine the influence of different input parameters on IPAR(0^-), K_1 , and K_2 , we carried out the correlation matrix analysis using the entire dataset. The performance of trained NN models is evaluated using the test dataset.

Figure 1 shows the correlation of IPAR(0^-), K_1 , and K_2 with atmospheric and oceanic parameters for open ocean waters. Figure 2 is the same as Fig. 1 except for coastal waters. It is obvious that the SZA and extinction AOD are significant influencers for surface IPAR for both ocean waters. For open ocean waters, K_1 and K_2 are mainly influenced by the SZA, [Chla], and AOD. For coastal waters, the absorption coefficient, backscattering coefficient, and backscattering fractions are also found to be major influential factors.

Figure 3 shows the density scatter plot between the NN network predicted values of IPAR(0^-), K_1 , and K_2 and the true values from the radiative transfer model for open ocean waters.

The solid line is the linear regression line with slope m and y intercept c . The R^2 values are all greater than 0.99, which shows the high correlation between predicted and true values. The RMSE and MAE values for the surface IPAR predictions are $11.93 \mu\text{mol photons m}^{-2} \text{s}^{-1}$ and $8.36 \mu\text{mol photons m}^{-2} \text{s}^{-1}$, respectively, which are very small numbers compared to the magnitude of IPAR that ranges from 11–1990 $\mu\text{mol photons m}^{-2} \text{s}^{-1}$ with majority values in the range of 1000–1900 $\mu\text{mol photons m}^{-2} \text{s}^{-1}$. Similar scenarios can be seen for the prediction of K_1 and K_2 . Furthermore, the slope of the linear regression fit is close to 1, and the bias c is close to 0 indicating great model performance.

Figure 4 shows the density scatter plot between the NN prediction and the true values for coastal waters. The dotted line indicates 1:1 line, and the solid line is the linear regression line with slope m and y intercept c . The R^2 values for surface IPAR, K_1 , and K_2 are high, i.e., 0.99, 0.97, and 0.98, respectively. The small RMSE, MAE, and y intercept show that the residual errors and bias for the model are very small. The NN performance in coastal waters is slightly worse than that of the open ocean due to the highly complex nature of waters represented by the large parameter space. As a limitation of NN models, it should be noted that NN models predicted small negative surface IPAR values for a small number of test cases (5 out of 1600) when the SZA is sufficiently large ($>82^\circ$) and the true surface IPAR values were less than $15 \mu\text{mol photons m}^{-2} \text{s}^{-1}$ for both water types. The performances of both NNs represented by statistical matrices are summarized in Table 3.

B. Discussion

The three predicted values IPAR(0^-), K_1 , and K_2 are used in Eq. (6) to give an accurate IPAR vertical profile up to euphotic zone depth. Using the IPAR vertical profile of each of the test

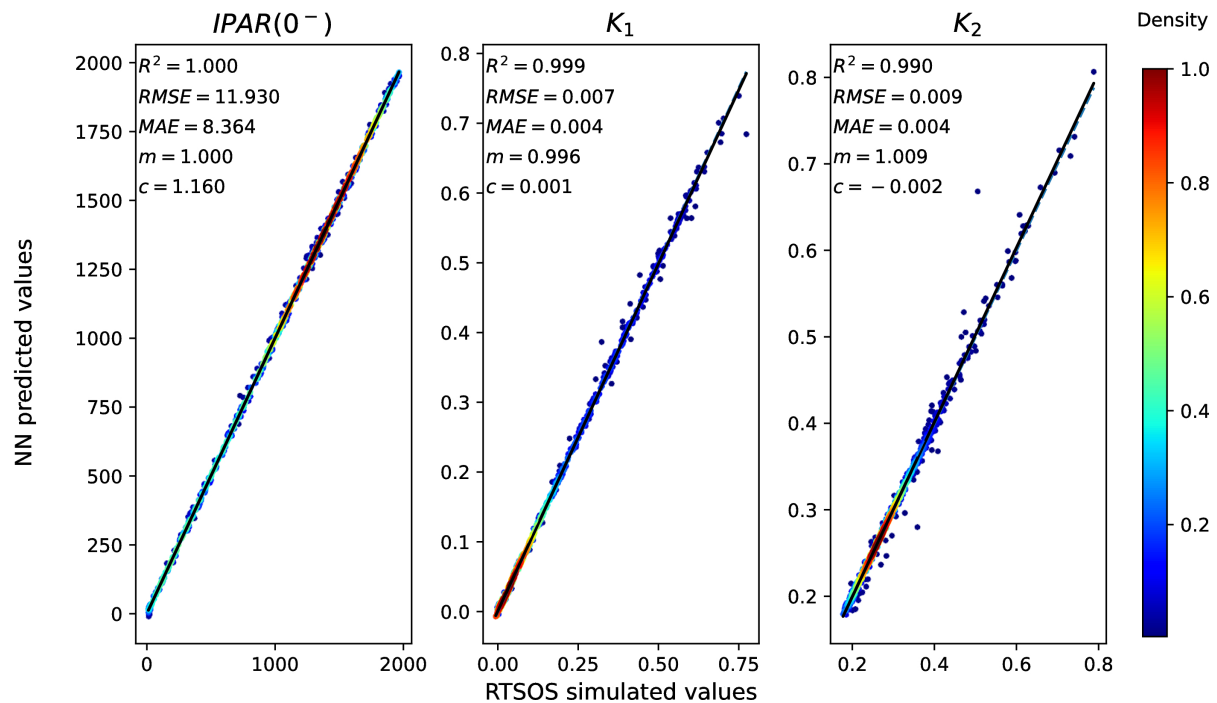


Fig. 3. Neural network predicted values of surface IPAR(0^-), K_1 , and K_2 for open ocean versus the true values from the RTSOS simulation.

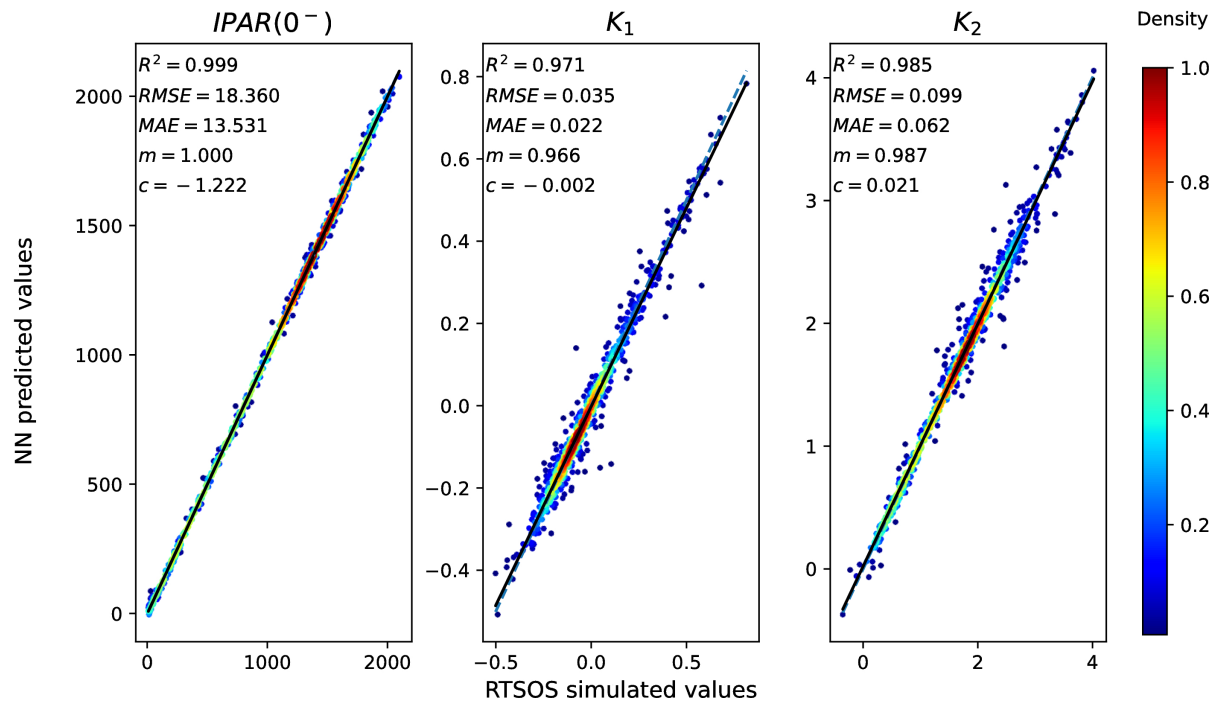


Fig. 4. Neural network predicted values of surface $IPAR(0^-)$, K_1 , and K_2 for coastal waters versus the true values from the RTSOS simulation.

Table 3. NN Models Performance in Open Ocean and Coastal Waters

NN Model	Predicted Variables	R^2	RMSE	MAE	m	c
Open Ocean	$IPAR(0^-)$	1	11.93	8.36	1	1.21
	K_1	0.99	0.007	0.004	0.99	0.001
	K_2	0.99	0.009	0.004	1	-0.002
Coastal Waters	$IPAR(0^-)$	0.99	18.36	13.53	1	-1.22
	K_1	0.97	0.03	0.02	0.96	-0.002
	K_2	0.98	0.09	0.06	0.98	0.02

cases, the diffuse attenuation coefficient at different depths is calculated using Eq. (12) and plotted against the true values simulated by RTSOS for open ocean (Fig. 5, left panel) and for coastal waters (Fig. 6, left panel). In Fig. 5, the NN predicts the diffuse attenuation coefficient at depths up to 100 m with RMSE of less than 0.011 m^{-1} and MAE of less than 0.01 m^{-1} for open ocean. In Fig. 6, the NN predicts the diffuse attenuation coefficient up to 20 m with an RMSE of less than 0.036 m^{-1} and MAE of less than 0.024 m^{-1} for coastal waters. The model performance below 100 m depth for open ocean and below 20 m depth for coastal waters is similar to that in shallow layers. The euphotic zone depths for open ocean cases ranged from 4 to 200 m with the majority less than 100 m and for coastal waters ranged from 2 to 176 m with the majority less than 25 m. These results confirm that our NN models can accurately fit the subsurface IPAR profile up to the euphotic zone depth.

The right panels of Figs. 5 and 6 show the density scatter plot between $K_{Lee}(z)$ and the true values from RTSOS simulation (K_{RTSOS}) for open ocean and coastal waters, respectively. The semi-analytical model works very well for the open ocean waters for the surface layer and has slight underprediction for deeper layers. The NN model has slightly high accuracy compared to

the semi-analytical model at deeper depths. However, for the coastal waters, the Lee model shows larger deviation from the true values. One reason for the Lee model's limited performance is the extensive parameter space of our dataset outside the range used by Lee for their model development. Another reason is the exclusion of the backscattering fraction parameter in their simulated dataset, which has been shown to have influence on the attenuation of IPAR below the surface for highly turbid coastal waters. In addition, we have used scalar irradiance in the calculation of IPAR as opposed to planar irradiance used in the Lee model. The authors of [39] have shown that there exists a difference of $>10\%$ between the diffuse attenuation coefficient calculated using planar irradiance and scalar irradiance. It should be noted that, for the accurate estimation of primary productivity, one should use the scalar irradiance to calculate IPAR [1].

It should be noted that values of A_{ph} obtained from [50] to calculate the absorption coefficient of phytoplankton are only applicable to ocean waters with [Chla] concentration up to $\sim 25 \text{ mgm}^{-3}$. For highly productive lake waters, where the [Chla] values can be significantly higher than 25, the spectral shape and magnitude of A_{ph} from [50] is found to deviate from the observed values [69]. While the dominant contributions

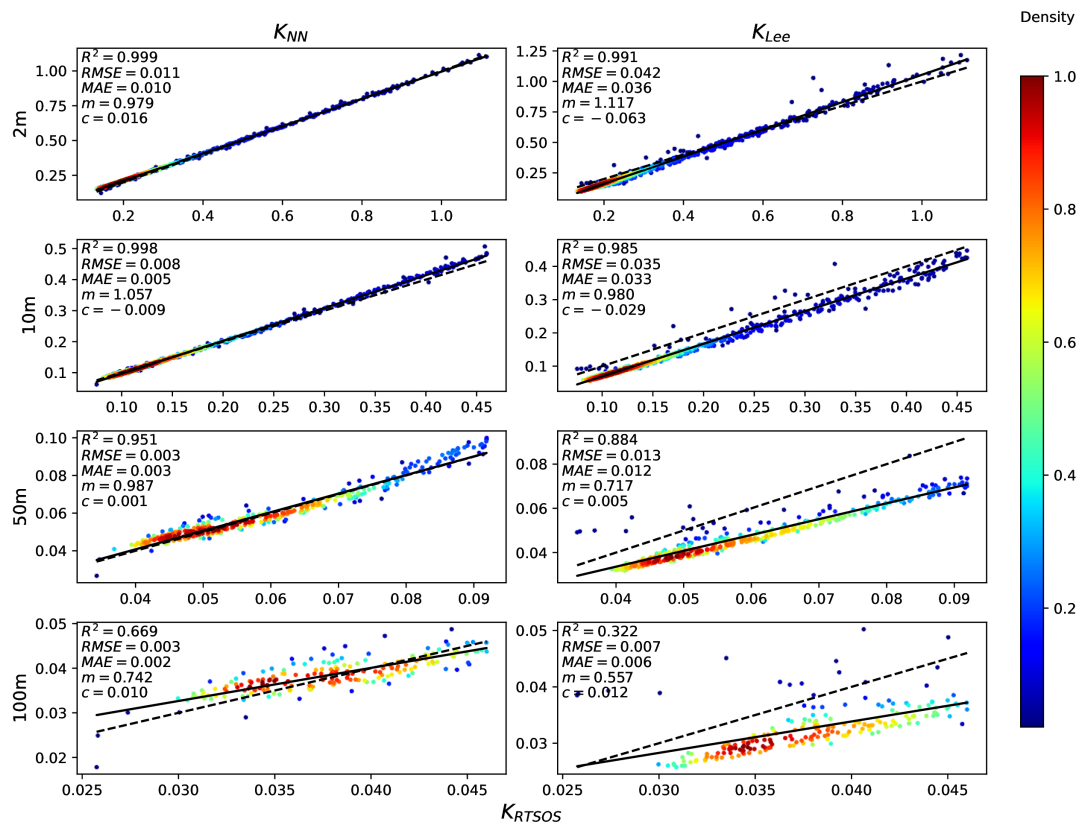


Fig. 5. Diffuse attenuation coefficients calculated from neural network predicted IPAR profile (left) and from Lee model (right) at different depth versus true values from the RTSOS simulation for open ocean waters.

modulating surface IPAR are the atmospheric and surface conditions, the impact of variations in spectral shape and magnitude of A_{ph} on the IPAR diffuse attenuation coefficients will be evaluated in a future study.

We have assumed a homogeneous distribution of IOPs below the ocean surface, which may cause poor performance of the NN models in waters with deep chlorophyll maxima (DCM) such as in Subtropical Gyre and Mediterranean Sea. This is true for other semi-analytical models as well [70]. The models are built for a clear sky, and the effect of cloud fractions in the IPAR calculation will be a subject of future studies. Meanwhile, the models can be combined with different cloud products to estimate IPAR in cloudy condition using the techniques as in [28].

1. Performance of Open Ocean Model on Coastal Waters Dataset and Vice Versa

Recognizing that it is not always possible to cleanly separate ocean observations into coastal and open ocean cases, we evaluated the performance of the IPAR models on test datasets for which they were not explicitly designed. Figure 7 shows the performance of the open ocean model on the coastal water dataset. The model predicts $IPAR(0^-)$ well, with $R^2 = 0.997$ and $RMSE = 55.67 \mu\text{mol photons m}^{-2} \text{s}^{-1}$, which is understandable as $IPAR(0^-)$ mostly depends on the atmospheric conditions and SZA. However, it cannot predict K_1 and K_2 , as the IOP ranges in coastal waters are out of the range for which

the open ocean model was trained. Figure 8 shows performance of the coastal waters model on the open ocean dataset. The coastal water IPAR model requires $a(490)$, $b_b(490)$, $B_p(670)$ as inputs, which were calculated by the same open ocean bio-optical models as those used in the radiative transfer simulation. The coastal waters model predicts $IPAR(0^-)$ and K_1 accurately with $R^2 > 0.98$. For K_2 , a small number of cases with small IPAR values corresponding to $\theta_0 > 82^\circ$ show larger deviations from the true value. The model provides a vertical profile of the diffuse attenuation coefficient with $RMSE < 0.032 \text{ m}^{-1}$. This RMSE value is small, albeit slightly larger than that of the open ocean model tested on same dataset ($RMSE \sim 0.01 \text{ m}^{-1}$). From this discussion, it can be concluded that the coastal waters model should be used in cases where the distinction between coastal and open ocean conditions cannot be well determined a priori.

4. CONCLUSION

Knowledge of IPAR at surface and its attenuation below the ocean surface is critical to understand many physical, biological, and chemical processes in the ocean body. A model that can compute the IPAR profile considering the influence of both atmosphere and ocean optical properties is important for accurate calculation of global marine net primary productivity, and the error contributed by the present semi-analytical models for IPAR profile is high in coastal regions.

To address this issue, we developed two IPAR prediction models based on NN, one each for open ocean and coastal

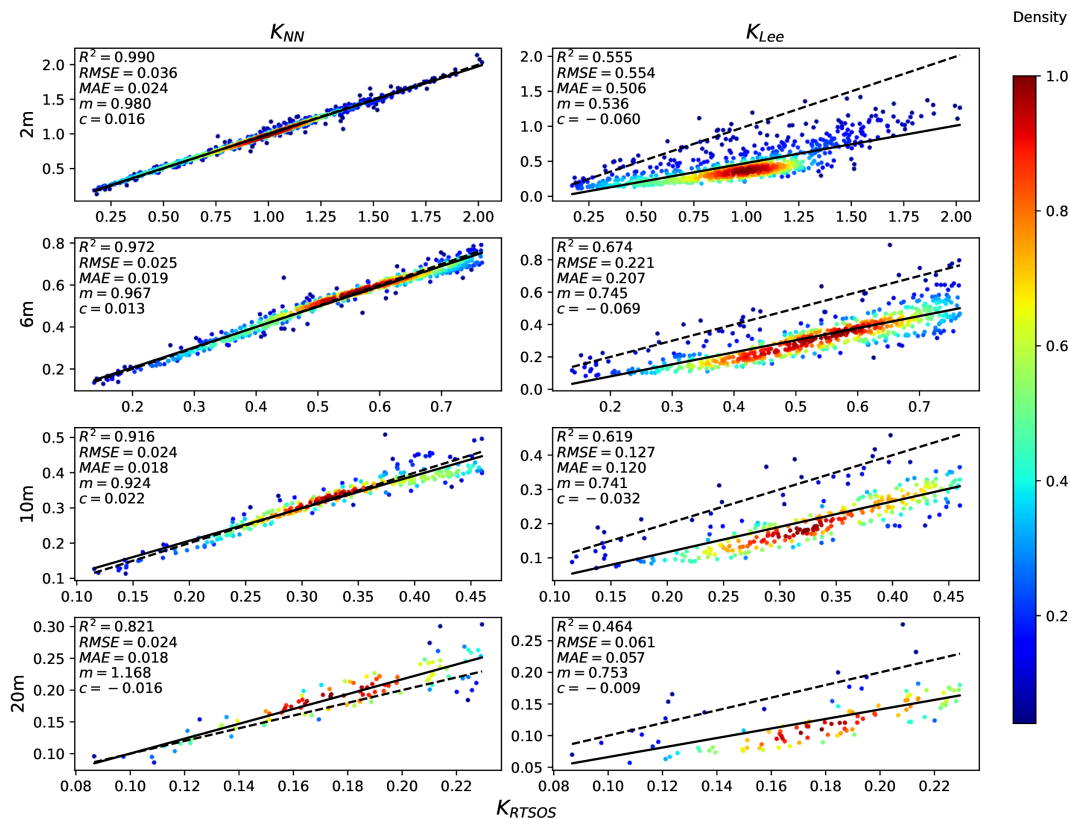


Fig. 6. Diffuse attenuation coefficients calculated from neural network predicted IPAR profile (left) and from Lee model (right) at different depth versus true values from the RTSOS simulation for coastal waters.

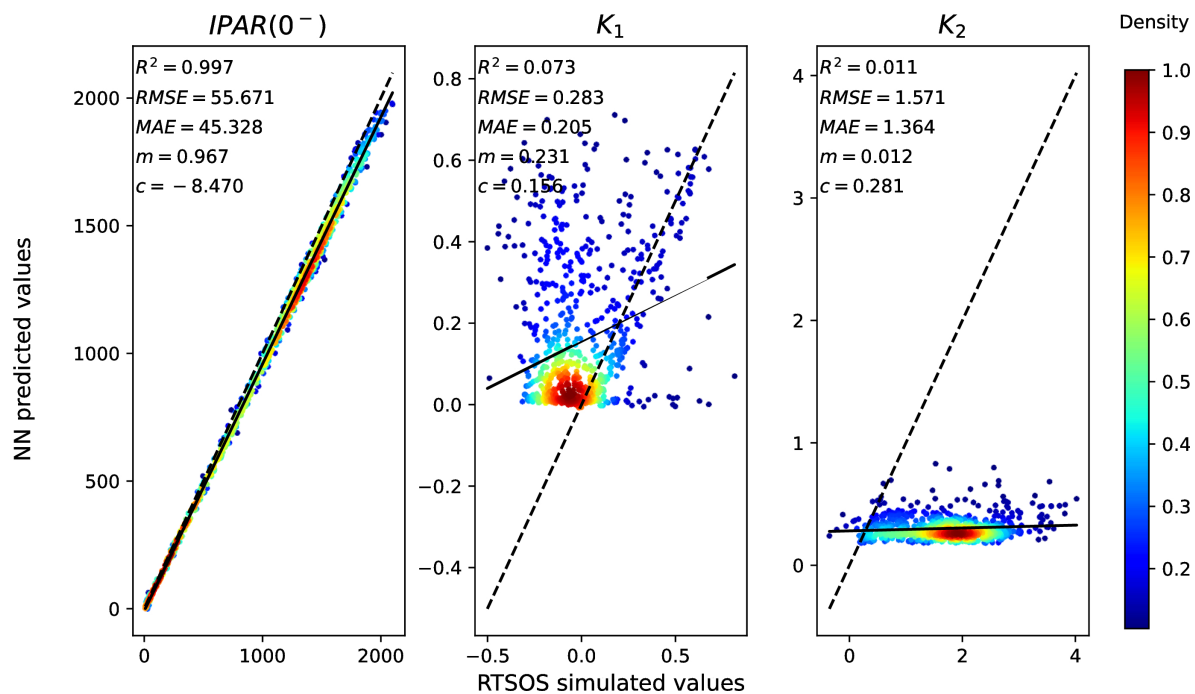


Fig. 7. Open ocean model's performance on coastal waters dataset.

waters. The models are trained and validated with a large volume of datasets covering a wide range of atmosphere and ocean conditions. Our NN models can predict surface IPAR ($IPAR(0^-)$), K_1 , and K_2 with RMSEs of 11.93 $\mu\text{mol photons}$

$\text{m}^{-2} \text{s}^{-1}$, 0.007 m^{-1} , and 0.009 $\text{m}^{-0.5}$, respectively, for open ocean waters and 18.36 $\mu\text{mol photons}$ $\text{m}^{-2} \text{s}^{-1}$, 0.03 m^{-1} , and 0.09 $\text{m}^{-0.5}$, respectively, for coastal waters. The R^2 values associated with $IPAR(0^-)$, K_1 , and K_2 predictions are greater

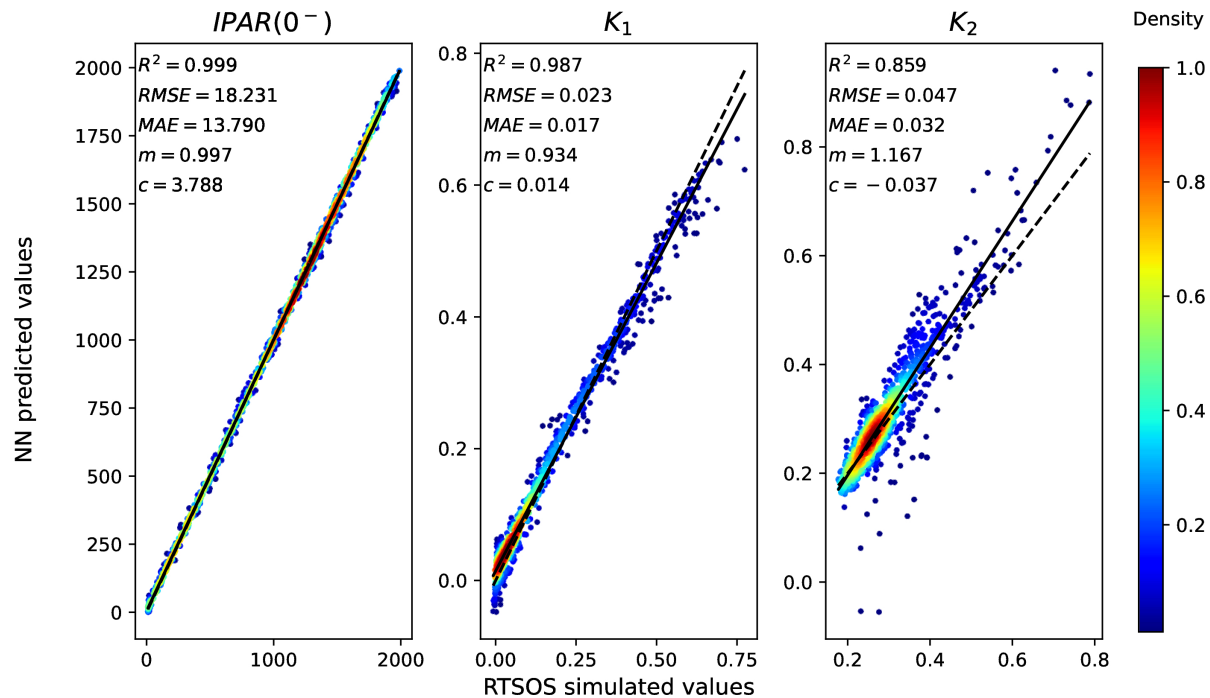


Fig. 8. Coastal waters model's performance on open ocean dataset.

than 0.99, 0.97, and 0.98, respectively, for both waters. The NN models can be used to calculate the diffuse attenuation coefficient of IPAR up to the euphotic zone with RMSE and MAE values less than 0.011 m^{-1} and 0.01 m^{-1} , respectively, in open ocean cases and 0.036 m^{-1} and 0.024 m^{-1} , respectively, in coastal waters cases. We also tested the performances of the IPAR models over different datasets, which suggests that the coastal waters NN model should be used in the areas where the distinction between coastal and open ocean is vague. These NN models are computationally more efficient than expensive radiative transfer models, yet they provide more accurate IPAR values than existing semi-analytical models in calculating IPAR and, hence, can facilitate the estimation of net primary production on global scales.

Funding. National Aeronautics and Space Administration (80NSSC20M0227).

Acknowledgment. The hardware used in the computational studies is part of the UMBC High Performance Computing Facility (HPCF). The facility is indirectly supported by the U.S. National Science Foundation through the MRI program (grant nos. CNS-0821258, CNS-1228778, and OAC-1726023) and the SCREMS program (grant no. DMS-0821311), with additional substantial support from the University of Maryland, Baltimore County (UMBC). See hpcf.umbc.edu for more information on HPCF and the projects using its resources.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are available in Ref. [71].

REFERENCES

- C. D. Mobley and E. S. Boss, "Improved irradiances for use in ocean heating, primary production, and photo-oxidation calculations," *Appl. Opt.* **51**, 6549–6560 (2012).
- P. G. Falkowski, R. T. Barber, and V. Smetacek, "Biogeochemical controls and feedbacks on ocean primary production," *Science* **281**, 200–206 (1998).
- J. T. Kirk, *Light and Photosynthesis in Aquatic Ecosystems* (Cambridge University, 1994).
- S. Sathyendranath, T. Platt, C. M. Caverhill, R. E. Warnock, and M. R. Lewis, "Remote sensing of oceanic primary production: computations using a spectral model," *Deep Sea Res. Part A* **36**, 431–453 (1989).
- T. Platt, S. Sathyendranath, C. M. Caverhill, and M. R. Lewis, "Ocean primary production and available light: further algorithms for remote sensing," *Deep-Sea Res. Part A* **35**, 855–879 (1988).
- B. Saulquin, A. Hamdi, F. Gohin, J. Populus, A. Mangin, and O. F. d'Andon, "Estimation of the diffuse attenuation coefficient K_{dPAR} using MERIS and application to seabed habitat mapping," *Remote Sens. Environ.* **128**, 224–233 (2013).
- M. J. Behrenfeld, R. T. O'Malley, E. S. Boss, T. K. Westberry, J. R. Graff, K. H. Halsey, A. J. Milligan, D. A. Siegel, and M. B. Brown, "Reevaluating ocean warming impacts on global phytoplankton," *Nat. Clim. Change* **6**, 323–330 (2016).
- M. J. Behrenfeld, T. K. Westberry, E. S. Boss, R. T. O'Malley, D. A. Siegel, J. D. Wiggert, B. A. Franz, C. R. McClain, G. C. Feldman, S. C. Doney, J. K. Moore, G. Dall'Olmo, A. J. Milligan, I. Lima, and N. Mahowald, "Satellite-detected fluorescence reveals global physiology of ocean phytoplankton," *Biogeosciences* **6**, 779–794 (2009).
- J. R. Morrison, "In situ determination of the quantum yield of phytoplankton chlorophyll a fluorescence: a simple algorithm, observations, and a model," *Limnol. Oceanogr.* **48**, 618–631 (2003).
- J. R. Morrison and D. S. Goodwin, "Phytoplankton photocompensation from space-based fluorescence measurements," *Geophys. Res. Lett.* **37**, L06603 (2010).
- P.-W. Zhai, E. Boss, B. Franz, P. J. Werdell, and Y. Hu, "Radiative transfer modeling of phytoplankton fluorescence quenching processes," *Remote Sens.* **10**, 1309 (2018).
- C. B. Field, M. J. Behrenfeld, J. T. Randerson, and P. Falkowski, "Primary production of the biosphere: integrating terrestrial and oceanic components," *Science* **281**, 237–240 (1998).

13. M. J. Behrenfeld and P. G. Falkowski, "Photosynthetic rates derived from satellite-based chlorophyll concentration," *Limnol. Oceanogr.* **42**, 1–20 (1997).
14. I. Asanuma, J. Nieke, K. Sasaoka, K. Matsumoto, and T. Kawano, "Optical properties control primary productivity model on the East China Sea," *Proc. SPIE* **4892**, 312–319 (2003).
15. T. Smyth, K. Pemberton, J. Aiken, and R. Geider, "A methodology to determine primary production and phytoplankton photosynthetic parameters from fast repetition rate fluorometry," *J. Plankton Res.* **26**, 1337–1350 (2004).
16. M. R. Lewis, M.-E. Carr, G. C. Feldman, W. Esaias, and C. McClain, "Influence of penetrating solar radiation on the heat budget of the equatorial pacific ocean," *Nature* **347**, 543–545 (1990).
17. A. Morel and D. Antoine, "Heating rate within the upper ocean in relation to its bio-optical state," *J. Phys. Oceanogr.* **24**, 1652–1665 (1994).
18. S. Sathyendranath, A. D. Gouveia, S. R. Shetye, P. Ravindran, and T. Platt, "Biological control of surface temperature in the Arabian sea," *Nature* **349**, 54–56 (1991).
19. Y. Wu, C. C. Tang, S. Sathyendranath, and T. Platt, "The impact of bio-optical heating on the properties of the upper ocean: a sensitivity study using a 3-D circulation model for the Labrador sea," *Deep Sea Res. Part II* **54**, 2630–2642 (2007).
20. R. W. Preisendorfer, "Secchi disk science: visual optics of natural waters 1," *Limnol. Oceanogr.* **31**, 909–926 (1986).
21. J. E. Tyler, "The Secchi disc," *Limnol. Oceanogr.* **13**, 1–6 (1968).
22. International Ocean Colour Coordinating Group (IOCCG), *Remote Sensing of Ocean Colour in Coastal, and Other Optically-Complex Waters*, S. Sathyendranath, ed. (2000), Vol. 3.
23. Z. Lee, *Remote Sensing of Inherent Optical Properties: Fundamentals, Tests of Algorithms, and Applications* (International Ocean Colour Coordinating Group, 2006).
24. K. L. Carder, F. R. Chen, Z. Lee, S. K. Hawes, and J. P. Cannizzaro, "MODIS ocean science team algorithm theoretical basis document," *ATBD* **19**, 7–18 (2003).
25. W. W. Gregg and K. L. Carder, "A simple spectral solar irradiance model for cloudless maritime atmospheres," *Limnol. Oceanogr.* **35**, 1657–1675 (1990).
26. P. Ricchiuzzi, S. Yang, C. Gautier, and D. Sowle, "SBDART: a research and teaching software tool for plane-parallel radiative transfer in the Earth's atmosphere," *Bull. Am. Meteorol. Soc.* **79**, 2101–2114 (1998).
27. R. Frouin, D. W. Lingner, C. Gautier, K. S. Baker, and R. C. Smith, "A simple analytical formula to compute clear sky total and photosynthetically available solar irradiance at the ocean surface," *J. Geophys. Res. Oceans* **94**, 9731–9742 (1989).
28. S. A. Somayajula, E. Devred, S. Bélanger, D. Antoine, V. Vellucci, and M. Babin, "Evaluation of sea-surface photosynthetically available radiation algorithms under various sky conditions and solar elevations," *Appl. Opt.* **57**, 3088–3105 (2018).
29. A. Morel, "Optical modeling of the upper ocean in relation to its biogenous matter content (case I waters)," *J. Geophys. Res. Oceans* **93**, 10749–10768 (1988).
30. A. Morel and S. Maritorena, "Bio-optical properties of oceanic waters: a reappraisal," *J. Geophys. Res. Oceans* **106**, 7163–7180 (2001).
31. A. Morel, Y. Huot, B. Gentili, P. J. Werdell, S. B. Hooker, and B. A. Franz, "Examining the consistency of products derived from various ocean color sensors in open ocean (case 1) waters in the perspective of a multi-sensor approach," *Remote Sens. Environ.* **111**, 69–88 (2007).
32. J. L. Mueller, "SeaWiFS algorithm for the diffuse attenuation coefficient, K (490), using water-leaving radiances at 490 and 555 nm," in *SeaWiFS Postlaunch Calibration Validation Analysis Part, Vol. 3*, (2000), pp. 24–27.
33. M. Wang, S. Son, and L. W. Harding, Jr., "Retrieval of diffuse attenuation coefficient in the Chesapeake Bay and turbid ocean regions for satellite ocean color applications," *J. Geophys. Res. Oceans* **114**, C10011 (2009).
34. Z. Lee, K. L. Carder, and R. A. Arnone, "Deriving inherent optical properties from water color: a multiband quasi-analytical algorithm for optically deep waters," *Appl. Opt.* **41**, 5755–5772 (2002).
35. Z. Lee, K. Du, R. Arnone, S. Liew, and B. Penta, "Penetration of solar radiation in the upper ocean: a numerical model for oceanic and coastal waters," *J. Geophys. Res. Oceans* **110**, C09019 (2005).
36. A. Albert and C. D. Mobley, "An analytical model for subsurface irradiance and remote sensing reflectance in deep and shallow case-2 waters," *Opt. Express* **11**, 2873–2890 (2003).
37. C. L. Gallegos, E. A. Lewis, and H.-C. Kim, *Coupling Suspended Sediment Dynamics and Light Penetration in the Upper Chesapeake Bay* (Smithsonian Environmental Research Center Edgewater, 2006).
38. P.-W. Zhai and Y. Hu, "An improved pseudo spherical shell algorithm for vector radiative transfer," *J. Quant. Spectrosc. Radiat. Transfer* **282**, 108132 (2022).
39. J. Wei and Z. Lee, "Model of the attenuation coefficient of daily photosynthetically available radiation in the upper ocean," *Methods Oceanogr.* **8**, 56–74 (2013).
40. P.-W. Zhai, Y. Hu, C. R. Trepte, and P. L. Lucker, "A vector radiative transfer model for coupled atmosphere and ocean systems based on successive order of scattering method," *Opt. Express* **17**, 2057–2079 (2009).
41. P.-W. Zhai, Y. Hu, J. Chowdhary, C. R. Trepte, P. L. Lucker, and D. B. Josset, "A vector radiative transfer model for coupled atmosphere and ocean systems with a rough interface," *J. Quant. Spectrosc. Radiat. Transfer* **111**, 1025–1040 (2010).
42. J. H. Halton, "On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals," *Numer. Math.* **2**, 84–90 (1960).
43. D. Loyola, M. Pedergrana, and S. G. Garca, "Smart sampling and incremental function learning for very large high dimensional data," *Neural Netw.* **78**, 75–87 (2016).
44. D. M. Drukker and R. Gates, "Generating halton sequences using mata," *Stata J.* **6**, 214–228 (2006).
45. P. Zhai, M. Gao, B. A. Franz, J. Werdell, A. Ibrahim, Y. Hu, and J. Chowdhary, "A radiative transfer simulator for pace: theory and applications," *Front. Remote Sens.* **3**, 840188 (2022).
46. R. M. Pope and E. S. Fry, "Absorption spectrum (380–700 nm) of pure water. II. Integrating cavity measurements," *Appl. Opt.* **36**, 8710–8723 (1997).
47. M. Twardowski, S. Freeman, S. Pegau, J. Zaneveld, J. Mueller, and E. Boss, *The Absorption Coefficient, an Overview* (NASA Ocean Optics Protocols, 2018), Chap. 1.
48. Z. Lee, J. Wei, K. Voss, M. Lewis, A. Bricaud, and Y. Huot, "Hyperspectral absorption coefficient of 'pure' seawater in the range of 350–550 nm inverted from remote sensing reflectance," *Appl. Opt.* **54**, 546–558 (2015).
49. J. D. Mason, M. T. Cone, and E. S. Fry, "Ultraviolet (250–550 nm) absorption spectrum of pure water," *Appl. Opt.* **55**, 7163–7172 (2016).
50. A. Bricaud, A. Morel, M. Babin, K. Allali, and H. Claustre, "Variations of light absorption by suspended particles with chlorophyll a concentration in oceanic (case 1) waters: analysis and implications for bio-optical models," *J. Geophys. Res. Oceans* **103**, 31033–31044 (1998).
51. C. D. Mobley, B. Gentili, H. R. Gordon, Z. Jin, G. W. Kattawar, A. Morel, P. Reinersman, K. Stamnes, and R. H. Stavn, "Comparison of numerical models for computing underwater light fields," *Appl. Opt.* **32**, 7484–7504 (1993).
52. G. R. Fournier and J. L. Forand, "Analytic phase function for ocean water," *Proc. SPIE* **2258**, 194–201 (1994).
53. K. J. Voss and E. S. Fry, "Measurement of the Mueller matrix for ocean water," *Appl. Opt.* **23**, 4427–4439 (1984).
54. C. Cox and W. Munk, "Measurement of the roughness of the sea surface from photographs of the sun's glitter," *J. Opt. Soc. Am.* **44**, 838–850 (1954).
55. E. P. Shettle and R. W. Fenn, "Models for the aerosols of the lower atmosphere and the effects of humidity variations on their optical properties," AFGL-TR 790214, U.S. Air Force Laboratory, Hanscom Air Force Base (1979), pp. 16.
56. Z. Ahmad, B. A. Franz, C. R. McClain, E. J. Kwiatkowska, J. Werdell, E. P. Shettle, and B. N. Holben, "New aerosol models for the retrieval of aerosol optical thickness and normalized water-leaving radiances from the SeaWiFS and MODIS sensors over coastal regions and open oceans," *Appl. Opt.* **49**, 5545–5560 (2010).

57. M. Gao, P.-W. Zhai, B. Franz, Y. Hu, K. Knobelspiesse, P. J. Werdell, A. Ibrahim, F. Xu, and B. Cairns, "Retrieval of aerosol properties and water-leaving reflectance from multi-angular polarimetric measurements over coastal waters," *Opt. Express* **26**, 8968–8989 (2018).
58. M. Gao, P.-W. Zhai, B. A. Franz, Y. Hu, K. Knobelspiesse, P. J. Werdell, A. Ibrahim, B. Cairns, and A. Chase, "Inversion of multiangular polarimetric measurements over open and coastal ocean waters: a joint retrieval algorithm for aerosol and water-leaving radiance properties," *Atmos. Meas. Tech.* **12**, 3921–3941 (2019).
59. L. Kou, D. Labrie, and P. Chylek, "Refractive indices of water and ice in the 0.65- to 2.5- μm spectral range," *Appl. Opt.* **32**, 3531–3540 (1993).
60. X. Zhang, L. Hu, and M.-X. He, "Scattering by pure seawater: effect of salinity," *Opt. Express* **17**, 5698–5710 (2009).
61. A. Morel, "Optical properties of pure water and pure sea water," *Opt. Aspects Oceanogr.* **1** (1974).
62. J. M. Sullivan and M. S. Twardowski, "Angular shape of the oceanic particulate volume scattering function in the backward direction," *Appl. Opt.* **48**, 6811–6819 (2009).
63. F. Chollet, "Keras," 2015, <https://keras.io>.
64. K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Netw.* **2**, 359–366 (1989).
65. J. A. Smith, "Lai inversion using a back-propagation neural network trained with a multiple scattering model," *IEEE Trans. Geosci. Remote Sens.* **31**, 1102–1106 (1993).
66. C. C. Aggarwal, *Neural Networks and Deep Learning* (Springer, 2018), Vol. **10**, pp. 978–983.
67. M. Gao, B. A. Franz, K. Knobelspiesse, P.-W. Zhai, V. Martins, S. Burton, B. Cairns, R. Ferrare, J. Gales, O. Hasekamp, Y. Hu, A. Ibrahim, B. McBride, A. Puthukkudy, P. J. Werdell, and X. Xu, "Efficient multi-angle polarimetric inversion of aerosols and ocean color powered by a deep neural network forward model," *Atmos. Meas. Tech.* **14**, 4083–4110 (2021).
68. D. Antanasijević, V. Pocajt, A. Perić-Grujić, and M. Ristić, "Modelling of dissolved oxygen in the Danube River using artificial neural networks and Monte Carlo simulation uncertainty analysis," *J. Hydrol.* **519**, 1895–1907 (2014).
69. B. Paavel, K. Kangro, H. Arst, A. Reinart, T. Kutser, and T. Noges, "Parameterization of chlorophyll-specific phytoplankton absorption coefficients for productive lake waters," *J. Limnol.* **75**, 423–438 (2016).
70. X. Xing and E. Boss, "Chlorophyll-based model to estimate underwater photosynthetically available radiation for modeling, in-situ, and remote-sensing applications," *Geophys. Res. Lett.* **48**, e2020GL092189 (2021).
71. K. Aryal, "NNIPAR neural network models," Python code, GitHub, 2021, <https://github.com/karyal1/NNIPAR>.