remote sensing

MDPI

Article

Enhanced Calculation of K_d (PAR) Using K_d (490) Based on a Recently Compiled Large In Situ and Satellite Database

Jorvin A. Zapata-Hinestroza [1,2], Eduardo Santamaría-del-Ángel [3,*], Alejandra Castillo-Ramírez [1], Sergio Cerdeira-Estrada [3], Adriana González-Silvera [1], Hansel Caballero-Aragón [3], Jesús A. Aguilar-Maldonado [4], Raúl Martell-Dubois [3], Laura Rosique-de-la-Cruz [3] and María-Teresa Sebastiá-Frasquet [4]

1 Facultad de Ciencias Marinas, Universidad Autónoma de Baja California, Ensenada 22860, Mexico; zapataj@uabc.edu.mx or yorvin1alexander@gmail.com (J.A.Z.-H.); alejandra.castillo.ramirez@uabc.edu.mx (A.C.-R.); adriana.gonzalez@uabc.edu.mx (A.G.-S.)
2 Sección de Litorales y Áreas Marinas, Dirección General Marítima, Santa Marta 470008, Colombia
3 Coordinación del Sistema de Información y Análisis Marino Costero, Comisión Nacional para el Conocimiento y Uso de la Biodiversidad, Mexico City 14010, Mexico; scerdeira@conabio.gob.mx (S.C.-E.); hcaballero@conabio.gob.mx or hansel.caballero@gmail.com (H.C.-A.); rmartell@conabio.gob.mx (R.M.-D.); laura.rosique@conabio.gob.mx or laura.rosique@ingenieria.unam.edu (L.R.-d.-l.-C.)
4 Institut d'Investigació per a la Gestió Integrada de Zones Costaneres (IGIC), Universitat Politècnica de València (UPV), Grau de Gandia, 46730 Gandia, Spain; jeagmal@upv.es or jesusaguilarmaldonado@gmail.com (J.A.A.-M.); mtsebastia@hma.upv.es (M.-T.S.-F.)
* Correspondence: santamaria@uabc.edu.mx or esantamaria@conabio.gob.mx

Highlights

What are the main findings?

- This paper presents a new approach for deriving $K_d\ (PAR)$ directly from the standard satellite $K_d\ (490)$ product, using a global database spanning diverse marine optical conditions. Among the tested methods, a power-law regression provided the best performance.
- The resulting model is accurate and robust, making it suitable for large-scale marine monitoring programs

What are the implications of the main finding?

- The best performing model is ready for operational implementation in monitoring programs to produce consistent $K_d\ (PAR)$ time series.
- This approach provides a reliable basis for studying water quality, phytoplankton dynamics, and other events affecting the light field in the water. It can be operationally estimated to support satellite monitoring programs and thus provide high-quality data for marine management programs.

Abstract

The vertical attenuation coefficient of photosynthetically active radiation ($K_d\ (PAR)$) is essential for characterizing the underwater light field and for operational marine monitoring. Although there have been efforts to use the standard satellite light attenuation product at 490 nm ($K_d\ (490)$) to estimate ($K_d\ (PAR)$) over a decade, earlier approaches were constrained by limited data. This study used a globally representative robust database of in-situ and satellite observations spanning diverse marine optical conditions and applied rigorous quality control. Three empirical models (linear, power, and a higher-order polynomial) were developed using four $K_d\ (490)$ satellite variants validated against an independent dataset and benchmarked against six published algorithms (36 total approximations). Performance was assessed using a Model Performance Index (MPI), where

check for updates

Academic Editor: Jorge Vazquez

Received: 4 October 2025
Revised: 17 November 2025
Accepted: 5 December 2025
Published: 10 December 2025

Citation: Zapata-Hinestroza, J.A.; Santamaría-del-Ángel, E.; Castillo-Ramírez, A.; Cerdeira-Estrada, S.; González-Silvera, A.; Caballero-Aragón, H.; Aguilar-Maldonado, J.A.; Martell-Dubois, R.; Rosique-de-la-Cruz, L.; Sebastiá-Frasquet, M.-T.. Enhanced Calculation of K_d (PAR) Using K_d (490) Based on a Recently Compiled Large In Situ and Satellite Database. *Remote Sens.* **2025**, *17*, 3990. https://doi.org/10.3390/rs17243990

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

values closer to 1 indicate a better model. The best model was a power regression driven by the standard satellite $K_d(490)$, which yielded an MPI of 0.8704, indicating a robust performance under a wide variability of marine optical conditions. These results highlight the value of multisensor products, which with a rigorous quality control protocol, could be used to estimate the $K_d\ (PAR)$ from the standard satellite $K_d\ (490)$. The objective of the proposed algorithm is to generate long-term $K_d\ (PAR)$ time series. This algorithm will be operational for implementation in marine ecosystem monitoring systems and can contribute to strengthening decision-making.

Keywords: vertical attenuation coefficient; photosynthetically available radiation; ocean color remote sensing; empirical algorithm

1. Introduction

Currently, assessing the impact of anthropogenic processes on ecosystem services (ES) and their associated habitats is a key issue in government sustainable development agendas. The UN 2030 Agenda for Sustainable Development [1] highlights the need and importance of monitoring and managing marine ecosystems through Sustainable Development Goals (SDGs) 13 (climate action) and 14 (marine life). In this context, the ocean is home to a wide variety of ES, making marine monitoring an essential tool for the planning and management of these ecosystems [2].

One of the primary anthropogenic impacts on the marine environment is the discharge of waste [3,4]. The introduction of new materials or chemicals alters the light field and optical properties of water, which manifests as a color change [5]. This optical variability can be quantified through the vertical diffuse light attenuation coefficient (K_d) [6], which depends on the incidence geometry of solar irradiance on the sea surface and the materials/chemicals in the water [7,8].

K_d is commonly evaluated in the visible range of the electromagnetic spectrum (400–700 nm), known as Photosynthetically Active Radiation (PAR), resulting in $K_d(PAR)$ [9,10]. Although $K_d(PAR)$ facilitates analyses by integrating the complete spectral range of light used by photosynthetic organisms, it is essential to stress that there is a spectral dependence of light attenuation. In general, wavelengths greater than 600 nm (red) are attenuated in the surface layers of the ocean, while those below 500 nm (blue) penetrate deeper [6,11]. This spectral variability is crucial for understanding the ecological and bio-optical processes that occur in different types of water.

$K_d(PAR)$ can be estimated directly using radiometers or indirectly with tools such as the Secchi Disk [10], and, more recently, satellite remote sensing [12]. This technique is based on the estimation of K_d at 490 nm ($K_d(490)$), a wavelength representative of the penetration of blue light across the water column, which penetrates to greater depths [13].

From satellite information, empirical algorithms have been developed that allow estimating $K_d(PAR)$ from $K_d(490)$ [14–18]. However, these approaches were developed more than a decade ago, when the availability of data generated with field multispectral spectroradiometers was still limited. Currently, the availability of hyper- and multispectral instruments (i.e., TriOS-RAMSES) has enabled a more detailed spectral characterization of the marine optical conditions. This technological advancement allows us to optimize $K_d(PAR)$ estimates based on $K_d(490)$ as current databases contain a greater number of observations and cover more areas and optical conditions than previous databases.

In addition, it is essential to consider that the relationship between $K_d(PAR)$ and $K_d(490)$ can vary significantly depending on the optical type of water [13,19,20]. In Case 1

waters, where optical properties are driven by seawater and phytoplankton (i.e., oceanic areas), this relationship could be clear. In contrast, in optically complex waters (i.e., coastal areas), where non-algal particles and dissolved organic matter are present, the relationship could become nonlinear, making it difficult to obtain direct empirical estimates. This variability highlights the importance of validating models with data that reflect this diversity.

In this context, the objective of this study is to enhance the estimation of $K_d(PAR)$ from $K_d(490)$ by developing validated empirical models using a robust, diverse, and up-to-date database representing different optical types of water on a global scale. This approximation aims to generate more accurate and reliable tools that can be integrated into marine monitoring systems supported by satellite remote sensors.

2. Materials and Methods

Data selection to estimate $K_d(PAR)$ from $K_d(490)$ was based on two primary criteria. First, the data shall represent the global variability of both parameters, avoiding dependence on a limited number of campaigns or a single region. Second, the datasets used to establish in situ and satellite relationships shall be of high quality and include a sufficient number of observations to develop and validate robust empirical models.

2.1. Databases

To ensure the representativeness and quality of the estimates, various databases covering a wide range of temporal and spatial data were used. These included the following: SeaWiFS Bio-optical Archive and Storage System (SeaBASS) [21] (1987–2017); NASA bio-Optical Marine Algorithm Dataset (NOMAD Version 2.0) [22] (1995–2007); Tara Ocean Consortium [23] (2009–2012); Biogeochemical-Argo data set (BGC-Argo full dataset version 2) [24] (1988–2023), and Phytoplankton Ecology Team (POPEYE) database of the Autonomous University of Baja California (UABC by its Spanish acronym) (2012–2024).

All the collected data (n = 18,208) underwent a quality control process to ensure the reliability of the estimates. Those stations that exhibited any of the following conditions were excluded: (1) profiles lacking surface records (<5 m) of downwelling irradiance just below the sea surface (E_{0-}); (2) profiles with negative irradiance values; (3) $K_d(PAR)$ and $K_d(490)$ values lower than the K_d of optically pure seawater (K_{dw}= 0.016 m^1; [12]); (4) measurements recorded on days with a high percentage (>30%) of cloud cover; (5) presence of whitecaps on the sea surface; and (6) measurements recorded when the solar angle was less than 45° relative to the sea surface. The criteria (4) to (6) followed the recommendations of Preisendorfer [25] and Castillo-Ramírez et al. [10,26]. The quality-control process produced a final set of 11,133 stations with in-situ $K_d(PAR)$ (hereafter $K_d(PAR)_{in\ situ}$) data (Figure 1 and Appendix A), 9710 of which have in-situ $K_d(490)$ (hereafter $K_d(490)_{in\ situ}$) data.

2.1.1. Considerations for In Situ Data

The $K_d(PAR)_{in\ situ}$ data were collected using one of the following four methods:

(1) The first method was when the data was already published in the consulted database. In this case these were used directly.

(2) The second method was where the stations had data on a downwelling irradiance PAR ($E_d(PAR)$), profile. In these cases, the Equation (1) [6] was used

$$ln\ (E_d(PAR)) = b_0 + (K_d(PAR) \times Z) \quad (1)$$

where the dependent variable is the natural logarithm of $E_d(PAR)$ and the independent variable Z is the depth at which such irradiance was measured.

(3) The third method was when the station had a profile of spectral downwelling irradiance ($E_d(\lambda)$). In this case to obtain $E_d(PAR)$ the $E_d(\lambda)$, between 400 and 700 nm were integrated; then, Equation (1) was applied.
(4) The fourth method was when the station reported Secchi disk depth readings (Z_{SD}). In these cases, were used the approaches reports by Castillo-Ramírez et al. [10].

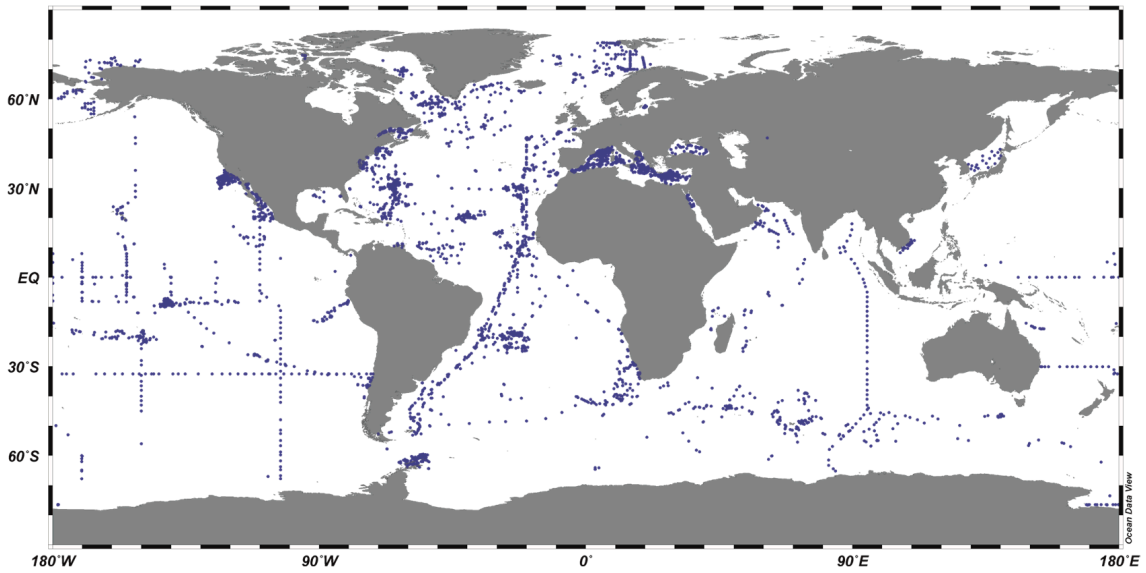


Figure 1. Global distribution of stations with $K_d(PAR)_{in\ situ}$ data.

Similarly, the $K_d(490)_{in\ situ}$ data were collected using one of the following two methods:

(1) The first method was when the $K_d(490)_{in\ situ}$ data was already published in the consulted database. In this case these were used directly.
(2) The second method was when the station had a profile of spectral downwelling irradiance ($E_d(\lambda)$). In this case the $E_d(490)$ profiles were selected. If the exactly the 490 nm profile was not available, the nearest value was used, provided it did not exceed a difference of ±10 nm. $K_d(490)_{in\ situ}$ was estimated using Equation (1) [6].

2.1.2. Considerations for Satellite Data

$K_d(490)$ based on satellite data (hereby $K_d(490)^{NASA/ESA}_{Standard}$) was calculated from the observations of the satellite platforms listed in Table 1. To this end, daily level 3 (L3) images with a 4 km spatial resolution per pixel were used, applying the standard product algorithm for each sensor [22]. It should be noted that all the images used in this study were captured at 4 km, including the OrbView-2 SeaWiFS images, which were processed by the POPEYE laboratory, from NASA Merged Local Area Coverage (MLAC) level 2 (L2) data. For the remaining sensors, L3 images were downloaded from the Ocean Color site [27]. With these images, daily multisensor composites were generated in accordance with Kahru et al. [28,29].

2.2. Numerical Procedures

This procedure resulted in a total of 14,339 stations with $K_d(490)^{NASA/ESA}_{Standard}$ data, which were extracted from the pixel values centered on the coordinates of the stations, following Santamaría-del-Ángel et al. [30]. Of these, 7788 have $K_d(PAR)_{in\ situ}$ data and 9238 $K_d(490)_{in\ situ}$ data.

Table 1. Satellite platforms, sensors, and inspections used for downloading $K_d(490)^{NASA/ESA}_{Standard}$.

Platform	Sensor	Revision Used
OrbView-2	Sea-viewing Wide Field-of-view (SeaWiFS)	SeaWiFS_R2022.0
TERRA	Moderate-Resolution Imaging Spectroradiometer (MODIS)	MODIST_R2022.0
ENVISAT	Medium Resolution Imaging Spectrometer (MERIS)	MERIS_R2022.0
AQUA	Moderate-Resolution Imaging Spectroradiometer (MODIS)	MODISA_R2022.0
Soumi-NPP	Visible Infrared Imaging Radiometer Suite (VIIRS)	VIIRS-SNPP_R2022.0
Sentinel-3A	Ocean and Land Colour Instrument (OLCI)	OLCIA-WRR ver. 003
NOAA-20	Visible Infrared Imaging Radiometer Suite (VIIRS)	VIIRS-JPSS1_R2022.0
Sentinel-3B	Ocean and Land Colour Instrument (OLCI)	OLCIB-WRR_ver. 003
NOAA-21	Visible Infrared Imaging Radiometer Suite (VIIRS)	VIIRS-JPSS2_R2022.0

Match-up analyses were applied according to Santamaria-del-Ángel et al. [30] for the following comparisons: $K_d(PAR)_{in\ situ}$ vs. $K_d(490)_{in\ situ}$; $K_d(PAR)_{in\ situ}$ vs. $K_d(490)^{NASA/ESA}_{Standard}$; and $K_d(490)_{in\ situ}$ vs. $K_d(490)^{NASA/ESA}_{Standard}$. The relationship between the data sets was evaluated by calculating Pearson's correlation coefficient ($r_{Pearson}$) (Equation (2)):

$$r_{Pearson} = \frac{Cov_{a,b}}{StDev_a \times StDev_b} \quad (2)$$

where $Cov_{a,b}$ is the covariance of the a and b, and $StDev_a$ and $StDev_b$ are the standard deviations of a and b, respectively. This coefficient is a measure of the linear association between two variables; it ranges between -1 and +1,where +1 indicates a direct linear relationship, -1 an inverse linear relationship, and zero indicates a nonlinear relationship.

Since $r_{Pearson}$ considers only linear relationships, Spearman's correlation coefficient was also used ($r_{Spearman}$) (Equation (3)) to evaluate probably nonlinear relationships. It is expressed by the covariance of the two variables divided by the product of their standard deviations.

$$r_{Spearman} = \frac{Cov_{Ra,Rb}}{StDev_{Ra} \times StDev_{Rb}} \quad (3)$$

where $Cov_{Ra,Rb}$ is covariance of the ranks of a and b; $StDev_{Ra}$ and $StDev_{Rb}$ is standard deviation of the ranks of a and b, respectively. The outcome for $r_{Spearman}$ is similar to that of $r_{Pearson}$, with a range of -1 to 1.

The statistical significance of the indices was evaluated [31], and the results were interpreted following Mu et al. [32]. Tables of critical values were obtained from Farnsworth and Triola [33] for Pearson's correlation coefficient and from [34] for Spearman's correlation coefficient.

2.3. Empirical Models

This study is based on empirical relationships between two variables, where $K_d(PAR)$ will always be the dependent variable and $K_d(490)$, the independent variable.

In the databases used, there are two types of $K_d(490)$:$K_d(490)_{in\ situ}$ and $K_d(490)^{NASA/ESA}_{Standard}$, both already described in previous sections.

Recently, Begouen-Demeaux et al. [24] proposed a new version of $K_d(490)^{NASA/ESA}_{Standard}$, recalculating the coefficients involved in the standard algorithms. This new version was called $K_d(490)^{NASA/ESA}_{New}$, and 10,929 observations are available in the BGC-ARGOS Database. Based on these data, a first-order (grade-one) model was developed to transform $K_d(490)^{NASA/ESA}_{Standard}$ data into $K_d(490)^{NASA/ESA}_{New}$ (Equation (4)).

$$K_d(490)^{NASA/ESA}_{New} = b_0 + (b_1 \times K_d(490)^{NASA/ESA}_{Standard}) \quad (4)$$

Separately, by observing a strong relationship between $K_d(490)_{in\ situ}$ and $K_d(PAR)_{in\ situ}$, $K_d(490)^{NASA/ESA}_{Standard}$ and $K_d(490)^{NASA/ESA}_{New}$ were corrected based on $K_d(490)_{in\ situ}$. To this end, three approximations were applied. The first is based on a linear model (Equations (5) and (6)); the second, on a power model (Equations (7) and (8)); and the third, on a Power Model based on a 4th-grade Taylor Polynomial, which will be called NASA-ESA Inspired Ocean Color Model (NESA-IOCM, hereby NESA) (Equations (9) and (10)).

$$K_d(490)_{\underline{\underline{NASA-ESA/Standard}}_{Corr}} = b_0 + (b_1 \times K_d(490)_{in\ situ}) \quad (5)$$

$$K_d(490)_{\underline{\underline{NASA-ESA/New}}_{Corr}} = b_0 + (b_1 \times K_d(490)_{in\ situ}) \quad (6)$$

$$K_d(490)_{\underline{\underline{NASA-ESA/Standard}}_{Corr}} = c_0 \times K_d(490)_{in\ situ}{}^{c_1} \quad (7)$$

$$K_d(490)_{\underline{\underline{NASA-ESA/New}}_{Corr}} = c_0 \times K_d(490)_{in\ situ}{}^{c_1} \quad (8)$$

$$K_d(490)_{\underline{\underline{NASA-ESA/Standard}}_{Corr}} = 10^{(a_0+(a_1\times\mathfrak{R})+(a_2\times\mathfrak{R}^2)+(a_3\times\mathfrak{R}^3)+(a_4\times\mathfrak{R}^4))} \quad (9)$$

$$K_d(490)_{\underline{\underline{NASA-ESA/New}}_{Corr}} = 10^{(a_0+(a_1\times\mathfrak{R})+(a_2\times\mathfrak{R}^2)+(a_3\times\mathfrak{R}^3)+(a_4\times\mathfrak{R}^4))} \quad (10)$$

where $\mathfrak{R} = log_{10}(K_d(490)_{in\ situ})$.

Based on the above, four versions of $K_d(490)$ were used in this study, each as an independent variable, in three different approaches based on empirical approximations for calculating $K_d(PAR)$ as a dependent variable. This resulted in a total of 12 models. Linear models are represented by Equations (11)–(14), while power models correspond to Equations (15)–(18), and NESA models are described by a general equation (Equation (19)).

$$K_d(PAR) = b_0 + \left(b_1 \times K_d(490)^{NASA/ESA}_{Standard}\right) \quad (11)$$

$$K_d(PAR) = b_0 + b_1 K_d(490)^{NASA/ESA}_{New} \quad (12)$$

$$K_d(PAR) = b_0 + (b_1 \times K_d(490)_{\underline{\underline{NASA-ESA/Standard}}_{Corr}}) \quad (13)$$

$$K_d(PAR) = b_0 + (b_1 \times K_d(490)_{\underline{\underline{NASA-ESA/New}}_{Corr}}) \quad (14)$$

$$K_d(PAR) = c_0 \times K_d(490)^{NASA/ESA c_1}_{Standard} \quad (15)$$

$$K_d(PAR) = c_0 \times K_d(490)^{NASA/ESA c_1}_{New} \quad (16)$$

$$K_d(PAR) = c_0 \times K_d(490)_{\underline{\underline{NASA-ESA/Standard}}_{Corr}}{}^{c_1} \quad (17)$$

$$K_d(PAR) = c_0 \times K_d(490)_{\underline{\underline{NASA-ESA/New}}_{Corr}}{}^{c_1} \quad (18)$$

$$K_d(PAR) = 10^{(a_0+(a_1\times\mathfrak{R})+(a_2\times\mathfrak{R}^2)+(a_3\times\mathfrak{R}^3)+(a_4\times\mathfrak{R}^4))} \quad (19)$$

where $\mathfrak{R}$, depending on the version of $K_d(490)$, can be:

$$\mathfrak{R} = log_{10}\left(K_d(490)^{NASA/ESA}_{Standard}\right)$$

$$\Re = log_{10}\left(K_d(490)_{New}^{NASA/ESA}\right)$$

$$\Re = log_{10}(K_d(490)\,\underset{Corr}{^{NASA-ESA/Standard}})$$

$$\Re = log_{10}(K_d(490)\,\underset{Corr}{^{NASA-ESA/New}})$$

To generate the models, following the recommendations of IOCCG [35] and Luijken et al. [36], the database was randomly split into two groups (50% for modeling and the remaining 50% for validation). To reduce random error in data selection, 10 different datasets were used for modeling, along with their respective validation sets. Each of these 10 datasets was evaluated in the models described above.

Statistical Significance for the Generated Empirical Models

To detect high-noise data in each model, a residual analysis was carried out following Castillo-Ramírez et al. [26] on the data used for modeling.

Likewise, the statistically significant contribution of the independent variable in each model was determined by applying a significance test of the coefficient associated with this variable, based on a *t*-test (Equation (20)).

$$t_{cal} = \frac{b_k}{SE_{bk}} \tag{20}$$

where b_k is the coefficient associated with the independent variable, and SE_{bk} (Equation (21)) is the standard error of the coefficient (b_k) expressed in the following equation:

$$SE_{bk} = \sqrt{\frac{\frac{\sum(y_i-\hat{y})^2}{n-(k+1)}}{\sum x^2 - \frac{(\sum x)^2}{n}}} \tag{21}$$

where x represents the independent variable and y_i is the dependent variable, $\hat{y}$ are the model values, n is the total number of observations.

The overall significance of each model was tested using an *F*-test based on Equation (22), following Castillo-Ramirez et al. [26]:

$$F_{cal} = \frac{\frac{\sum(y_i-\overline{y})^2-\sum(y_i-\hat{y})^2}{k}}{\frac{\sum(y_i-\hat{y})^2}{n-(k+1)}} \tag{22}$$

where y_i is the dependent variable, $\overline{y}$ is the y_i mean, $\hat{y}$ are the model values, n is the total number of observations, and k is the number of independent variables.

The proportion of the variation of the dependent variable that can be explained by the independent variables was estimated with the coefficient of determination (R^2) (Equation (23)).

$$R^2 = \left(\frac{\sum(y_i-\hat{y})^2}{\sum(y_i-\overline{y})^2}\right) \times 100 \tag{23}$$

2.4. Model Validation

For validation process, the three models proposed in this study were each tested using the four available versions of $K_d(490)$, resulting in 12 approaches. Additionally, six models reported in the literature (Table 2) were also tested with the same four versions of $K_d(490)$, yielding 24 additional approaches. In total, 36 approaches were evaluated.

Table 2. Models reported in the literature to estimate $K_d(PAR)$ from $K_d(490)$.

Reference	Model
Slope-Weighting Model derived from a first-degree, first-order polynomial (SWM) (Kratzer et al. [37]; Tang et al. [38]).	$K_d(PAR) = b_{1in\ situ} \times K_d(490)$ where $b_{1in\ situ}$ $K_d(PAR) = b_0 + (b_{1in\ situ} \times K_d(490)_{in\ situ})$
Morel et al. [14]	If $K_d(490) \leq 0.3\ m^{-1}$ $K_d(PAR) = 0.0864 + (0.884 \times K_d(490)) - (0.00137 \times [K_d(490)]^{-1})$ If $K_d(490) > 0.3\ m^{-1}$ $K_d(PAR) = 0.0665 + (0.874 \times K_d(490)) - (0.00121 \times [K_d(490)]^{-1})$
Pierson et al. [15]	$K_d(PAR) = 0.1134 + (0.6098 \times K_d(490))$
Pierson et al. [15]	$K_d(PAR) = 0.6677 \times K_d(490)^{0.6763}$
Wang et al. [16]	$K_d(PAR) = 0.8045 \times K_d(490)^{0.917}$
Saulquin et al. [18]	If $K_d(490) \leq 0.115\ m^{-1}$ $K_d(PAR) = (4.6051 \times K_d(490))/((6.0700 \times K_d(490)) + 3.200)$ If $K_d(490) > 0.115\ m^{-1}$ $K_d(PAR) = 0.8100 \times K_d(490)^{0.8256}$

This was carried out with 50% of the data assigned for validation, yielding an observed value ($K_d(PAR)_{in\ situ}$) and a modeled value ($K_d(PAR)_{model}$) for each model. From these values, three statistical descriptors were calculated: Root Mean Square Deviation ($RMSD$) (Equation (24)), Bias Analysis ($BIAS$) (Equation(25)) and Mean Absolute Percentage Error (MAPE) (Equation (26)) and Integrated Absolute Residuals (Equation (27)).

$$RMSD = \sqrt{\frac{\sum(K_d(PAR)_{in\ situ} - K_d(PAR)_{model})^2}{n}} \quad (24)$$

$$BIAS = Mean\ (K_d(PAR)_{in\ situ} - K_d(PAR)_{model}) \quad (25)$$

$$MAPE = \frac{\sum \frac{|K_d(PAR)_{in\ situ} - K_d(PAR)_{model}\ |\ \times 100}{K_d(PAR)_{in\ situ}}}{n} \quad (26)$$

where n is the total number of data in the analysis, $K_d(PAR)_{in\ situ} - K_d(PAR)_{model}$ are the residuals, and $|K_d(PAR)_{in\ situ} - K_d(PAR)_{model}|$ is the absolute value of the residuals. Likewise, $BIAS$ is the mean difference of the residuals. In $RMSD$, $BIAS$, and $MAPE$, the values closest to zero represent the best results [10,39,40].

To determine the best approach when comparing several models based on the same dependent and independent variables, the Model Performance Index (MPI) can be used [10,26]. Originally, this was based on the Mean Absolute Error (MAE), $RMSD$, and $BIAS$; however, in this study, MAE was replaced by $MAPE$ (Equation (27)), as both parameters express the absolute mean of the residuals, but $MAPE$ weights the residuals according to the observed value, making it more sensitive.

$$MPI = 1 - \left(\frac{\left(\frac{R_{RMSD}}{p}\right) + \left(\frac{R_{|BIAS|}}{p}\right) + \left(\frac{R_{MAPE}}{p}\right)}{3} \right) \quad (27)$$

where R_{RMSD} and R_{MAPE} are the individual ranks for $RMSD$ and $MAPE$, respectively. $R_{|BIAS|}$ is the individual absolute ranks of $BIAS$, and p is the total number of models compared. Ranks were calculated following Wilcoxon [41]. MPI ranges between 0 and 1, where the values closest to 1 indicate better model performance.

Once the best model was determined, and with the validation data, a comparison was made based on matchup charts [30]. This graphical analysis was reinforced with $RMSD$, and $BIAS$, and Integrated Absolute Residuals (IAR) (Equation (28))

$$IAR = \sum |(K_d(PAR)_{in\ situ} - K_d(PAR)_{model})| \quad (28)$$

Low values are associated with a better model, while higher values indicate a less accurate model.

3. Results

3.1. Database

The consulted databases initially yielded 18,208 $K_d(PAR)_{in\ situ}$. After applying a quality control protocol, the dataset was reduced to 11,161 records. Subsequently, data with $K_d(PAR)$ values > 1 were excluded, resulting in a final total of 11,133 stations (Figure 1). The classification of this data into water types by water (Figure 2) reveals an uneven distribution: Clear Waters ($K_d(PAR)$ < 0.02) account for 10,456 observations, Transitional Waters ($K_d(PAR)$ 0.2–0.4) for 534, and Turbid Waters ($K_d(PAR)$ > 0.4) for only 143.

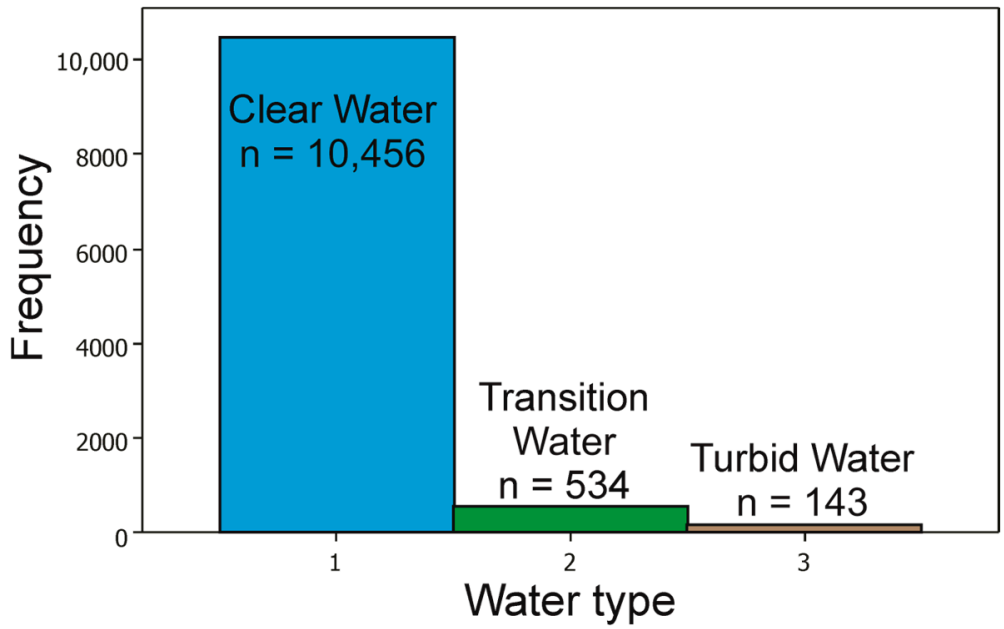


Figure 2. Histogram of water types derived from $K_d(PAR)_{in\ situ}$. Water types were described according with Castillo-Ramirez et al. [10].

Once the data were screened, scatter plots were constructed to compare the data obtained through in situ measurements and satellite estimates (Figure 3). In Figure 3a, the relationship between $K_d(PAR)_{in\ situ}$ and $K_d(490)_{in\ situ}$ shows a high correlation ($r_{Pearson} = 0.95$, $r_{Spearman} = 0.85$).

The Figure 3b shows the relationship between $K_d(PAR)_{in\ situ}$ and $K_d(490)^{NASA/ESA}_{Standard}$, where a greater dispersion and a lower correlation are observed ($r_{Pearson} = 0.73$; $r_{Spearman} = 0.80$). The comparison between $K_d(PAR)_{in\ situ}$ and $K_d(490)^{NASA/ESA}_{Standard}$ (Figure 4) yields higher correlation coefficients ($r_{Pearson} = 0.78$; $r_{Spearman} = 0.82$), than those observed in Figure 3b.

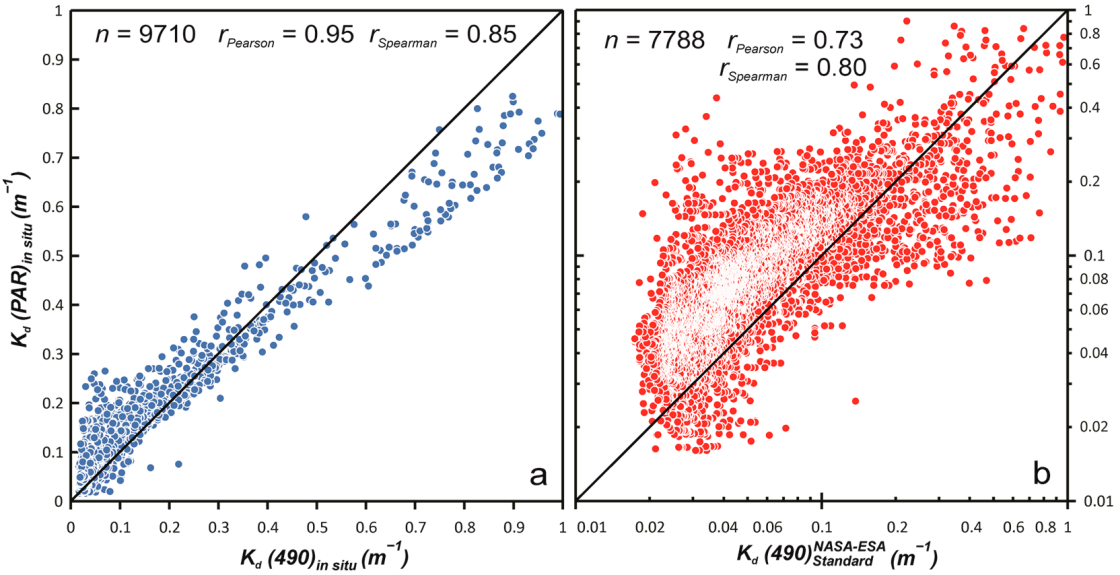


Figure 3. Scatter plots. (**a**) Comparison between $K_d(PAR)_{in\ situ}$ and $K_d(490)_{in\ situ}$. (**b**) Relationship between $K_d(PAR)_{in\ situ}$ and $K_d(490)^{NASA/ESA}_{Standard}$.

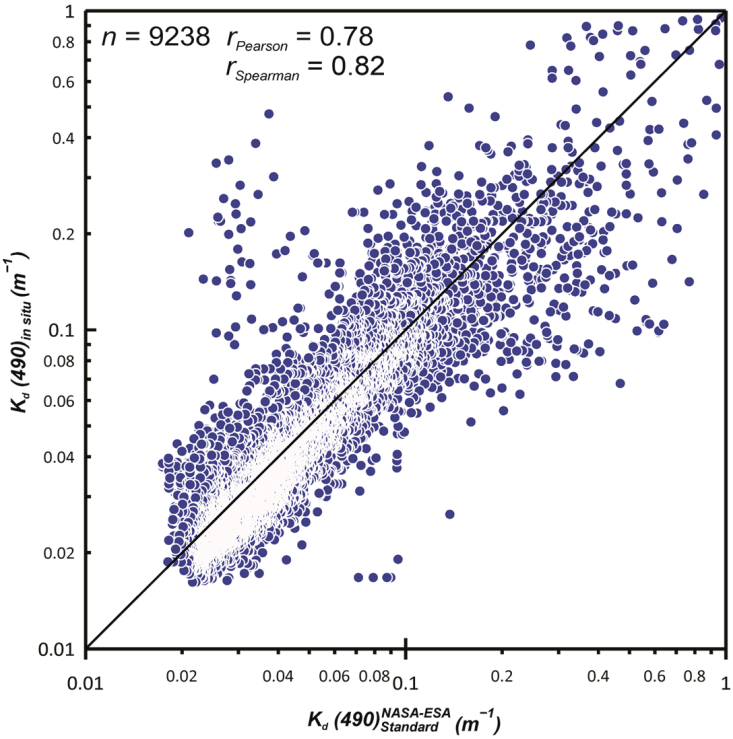


Figure 4. Scatter plot of the relationship between $K_d(490)_{in\ situ}$ and $K_d(490)^{NASA/ESA}_{Standard}$.

3.2. Models

3.2.1. $K_d(490)$ Versions

The Figure 3a, suggests the use of linear empirical models to estimate $K_d(PAR)_{in\ situ}$ from $K_d(490)_{in\ situ}$. From the available data, a first-order and first-degree polynomial was developed in which $K_d(PAR)_{in\ situ}$ was the dependent variable and $K_d(490)_{in\ situ}$ the independent variable (Equation (29));

$$K_d(PAR)_{in\ situ} = 0.0234 + (0.909 \times K_d(490)_{in\ situ}) \quad (29)$$

From this equation (Equation (29)), the slope coefficient (0.909) was used to develop a version of the global SWM, which could be evaluated and compared with the other empirical models reported in the literature (Equation (30));

$$K_d(PAR) = 0.909 \times K_d(490) \quad (30)$$

where $K_d(490)$ would be each of the versions developed in the present study.

The A new corrected version of $K_d(490)^{NASA/ESA}_{Standard}$, was develop ($K_d(490)^{NASA/ESA}_{New}$), using the BC-ARGOS database. The linear relationship between $K_d(490)^{NASA/ESA}_{Standard}$ and $K_d(490)^{NASA/ESA}_{New}$ (Figure 5). This relationship led to the development of Equation (31).

$$K_d(490)^{NASA/ESA}_{New} = 0.003028 + (0.805 \times K_d(490)^{NASA/ESA}_{Standard}) \quad (31)$$

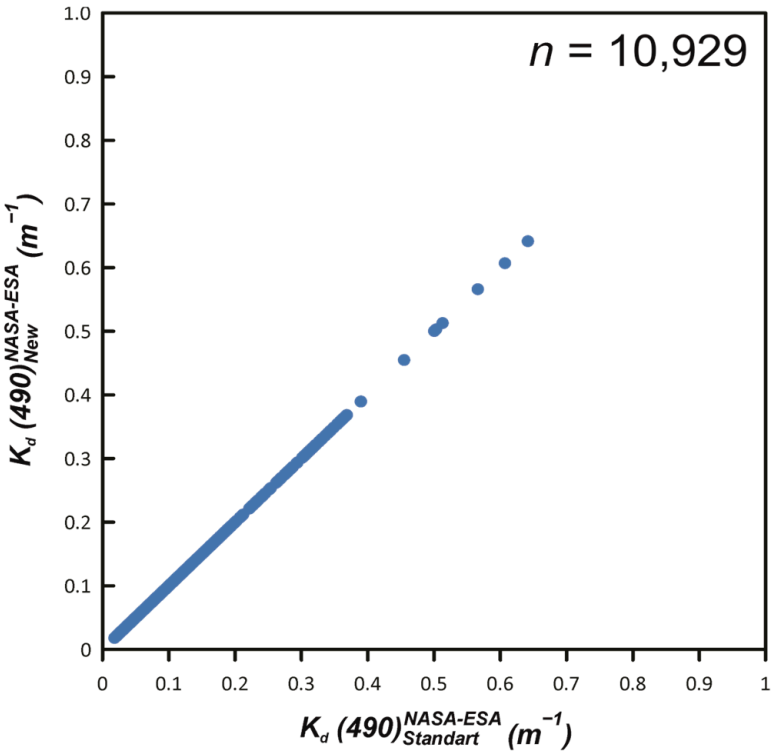


Figure 5. Relationship between $K_d(490)^{NASA/ESA}_{Standard}$ and $K_d(490)^{NASA/ESA}_{New}$.

The model coefficients were statistically significant (intercept: $t_{crit} = 1.96$, $t_{calc} = 256.22$; slope: $t_{calc} = 3959.23$; $F_{crit} = 3.84$, $F_{calc} = 15{,}775{,}508.45$; $n = 10{,}929$; $\alpha = 5\%$), and $R^2 = 99.9$.

Based on the logic of the above-mentioned correction and on the relationship between $K_d(PAR)_{in\ situ}$ and $K_d(490)_{in\ situ}$ (Figure 3a), we explored the correction of $K_d(490)^{NASA/ESA}_{Standard}$ as if it were $K_d(490)_{in\ situ}$. To this end, different empirical approaches were evaluated, including the linear, power, and NESA models.

3.2.2. Models Validation

The models were validated based on the MPI (Equation (27)) (Table 3), and the results showed that the power model yielded the best approximations. Again, this suggests that the relationship between $K_d(490)^{NASA/ESA}_{Standard}$ (or $K_d(490)^{NASA/ESA}_{New}$) and $K_d(490)_{in\ situ}$ is not strictly linear, but follows a power function. The models with the best performance are shown in Equations (32) and (33).

Table 3. Statistical descriptors results for validating empirical models to estimate $K_d(490)_{\underline{Corr}}^{NASA-ESA/Standard}$ and $K_d(490)_{\underline{Corr}}^{NASA-ESA/New}$. IV, independent variable.

Type of Model	IV	RMSD	BIAS	MAPE	MPI
Linear	$K_d(490)_{Standard}^{NASA/ESA}$	0.0435	0.0030	23.3178	0.3333
Power		0.0428	0.0041	21.5977	0.4444
NESA		0.0453	0.0033	22.7503	0.2222
Linear	$K_d(490)_{New}^{NASA/ESA}$	0.0403	0.0015	23.0121	0.3333
Power		0.0397	0.0022	21.8091	0.4444
NESA		0.0437	0.0027	21.0465	0.2222

$$K_d(490)_{\underline{Corr}}^{NASA-ESA/Standard} = 0.669 \times K_d(490)_{in\ situ}^{0.87} \quad (32)$$

$$K_d(490)_{\underline{Corr}}^{NASA-ESA/New} = 0.934 \times K_d(490)_{in\ situ}^{0.938} \quad (33)$$

Based on the models developed (Equations (31)–(33)), we defined the four versions of $K_d(490)$ used as the independent variable in the three approaches developed (linear, power, and NESA) $K_d(490)_{Standard}^{NASA/ESA}$, $K_d(490)_{New}^{NASA/ESA}$, $K_d(490)_{\underline{Corr}}^{NASA-ESA/Standard}$, and $K_d(490)_{\underline{Corr}}^{NASA-ESA/New}$.

Considering these four versions of the independent variable ($K_d(490)$), the three types of models (linear, power, and NESA) were applied to each of them to model $K_d(PAR)$ as the dependent variable. To evaluate the effectiveness of each approach, the coefficients of each model, its statistical significance, and fitness metrics were calculated (Tables 4–6).

Table 4. Linear regression models to estimate $K_d(PAR)$ based on the four variants of $K_d(490)$. The test statistics for coefficient significance and the overall model significance are shown.

Linear Model $K_d(PAR) = b0 + (b_1 \times IV)$										
Equation	***n***	**IV**	R^2	b_0	t_{b0calc}	b_1	t_{b1calc}	t_{crit} $\alpha = 5\%$ $df = n-p$	F_{calc}	F_{crit}
(11)	3733	$K_d(490)_{Standard}^{NASA/ESA}$	72.2	0.0380	58.45	0.740	98.39	1.96	9681.32	2.99
(12)	3729	$K_d(490)_{New}^{NASA/ESA}$	72.1	0.0334	48.93	0.961	98.21	1.96	9645.68	2.99
(13)	3717	$K_d(490)_{\underline{Corr}}^{NASA-ESA/Standard}$	74.9	0.0245	35.39	1.03	105.34	1.96	11,097.47	2.99
(14)	3719	$K_d(490)_{\underline{Corr}}^{NASA-ESA/New}$	75.3	0.0273	40.72	0.970	106.57	1.96	11,356.68	2.99

Table 5. Power regression models to estimate $K_d(PAR)$ based on the four variants of $K_d(490)$. The test statistics for coefficient significance and the overall model significance are shown. IV, independent variable.

Power Regression $K_d(PAR) = c_0 \times IV^{c1}$										
Equation	***n***	**IV**	R^2	c_0	t_{b0calc}	c_1	t_{b1calc}	t_{crit} $\alpha = 5\%$ $df = n-p$	F_{calc}	F_{crit}
(15)	3804	$K_d(490)_{Standard}^{NASA/ESA}$	70.2	0.575	−24.72	0.683	94.53	1.96	8936.38	2.99
(16)	3799	$K_d(490)_{New}^{NASA/ESA}$	70.5	0.737	−12.39	0.732	95.22	1.96	9066.07	2.99
(17)	3797	$K_d(490)_{\underline{Corr}}^{NASA-ESA/Standard}$	72.0	0.807	−8.73	0.791	98.68	1.96	9738.64	2.99
(18)	3797	$K_d(490)_{\underline{Corr}}^{NASA-ESA/New}$	71.9	0.761	−11.07	0.775	98.44	1.96	9691.13	2.99

Table 6. NESA Models to estimate $K_d(PAR)$ based on the four variants of $K_d(490)$. The test statistics for coefficient significance and the overall model significance are shown. IV, independent variable.

NESA Model $K_d(PAR) = 10^{(a_0+(a_1\times\Re)+(a_2\times\Re^2)+(a_3\times\Re^3)+(a_4\times\Re^4))}$ where $\Re = log_{10}(IV)$															
Equation	***n***	**IV**	R^2	a_0	t_{calc}	a_1	t_{calc}	a_2	t_{calc}	a_3	t_{calc}	a_4	t_{calc}	t_{crit}	F_{calc}	F_{crit}
(19)	3805	$K_d(490)^{NASA/ESA}_{Standard}$	71.2	−0.17	−3.56	2.68	10.27	4.78	10.13	3.77	10.98	0.96	11.10	1.96	2344.83	2.37
(19)	3808	$K_d(490)^{NASA/ESA}_{New}$	71.1	0.04	0.55	3.36	9.05	5.59	9.19	4.09	9.96	0.99	10.06	1.96	2337.92	2.37
(19)	3792	$K_d(490)_{\frac{NASA-ESA/StandardD}{Corr}}$	72.7	0.72	5.76	5.92	10.18	9.30	9.90	6.54	10.32	1.58	10.38	1.96	2525.14	2.37
(19)	3792	$K_d(490)_{\frac{NASA-ESA/New}{Corr}}$	72.7	0.41	4.33	4.77	10.24	7.73	9.88	5.61	10.33	1.38	10.35	1.96	2525.77	2.37

The Tables 4–6 show that the coefficients of each model were statistically significant, as were all the approximations evaluated.

For validation, the MPI (Equation (27)) (Figure 6, Appendix B) was estimated for the three empirical models (linear, power, and NESA) and the six models reported in the literature (Table 2). The four versions of K_d (490) were applied to each of these models, resulting in a total of 36 validated approximations.

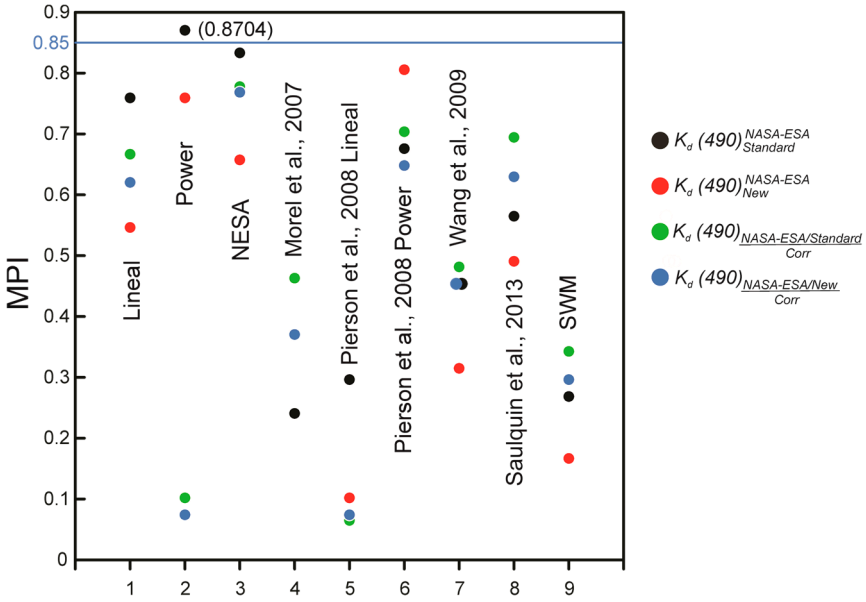


Figure 6. Scatterplot of the MPIs (y axis) for the 9 models (x axis) (three from this work and six reported in literature). The literature models are: Morel et al. [14]; Pierson et al. [15], linear and power versions; Wang et al., [16]; Saulquin et al. [18] and SWM [37,38]. Color dots indicate the K_d (490) model used as the independent variable. The position of the colored points on the x axis corresponds to each of the 9 contrasted models, while their height on the y axis corresponds to the MPI value described in Appendix B). The best model is indicated by placing the MPI value next to the symbol.

The evaluation of 36 approximations showed that the power model developed in this study, using $K_d(490)^{NASA/ESA}_{Standard}$ as the independent variable (Equation (15)), returned the best performance.

This Power model (Equation (15)) was compared with the six models reported in the literature (Table 2), using the same validation dataset described above and $K_d(490)^{NASA/ESA}_{Standard}$ as the independent variable. This comparison was based on matchup plots (Figure 7). All models showed similar point cloud shapes. The Pierson et al. [15] model, linear version (Figure 7c), is the one that deviates most from the 45-degree line. The Morel et al. [14] model (Figure 7b) showed an overestimation, while the Wang et al., [16] Saulquin et al. [18], and SWM [37,38] models (Figure 7e–g) showed an underestimation. The Power model and the Pierson et al. [15] model, Power version (Figure 7a,d), show a better distribution with

respect to the 1:1 line. Analyzing the $RMSD$(Equation (24)), $BIAS$ (Equation (25)) and IAR (Equation (28)) confirms that the Power Model presented in this work has the best fit when comparing $K_d(PAR)_{in\ situ}$ and $K_d(PAR)_{model}$, as it presented the lowest values among all the models in all these statistical descriptors.

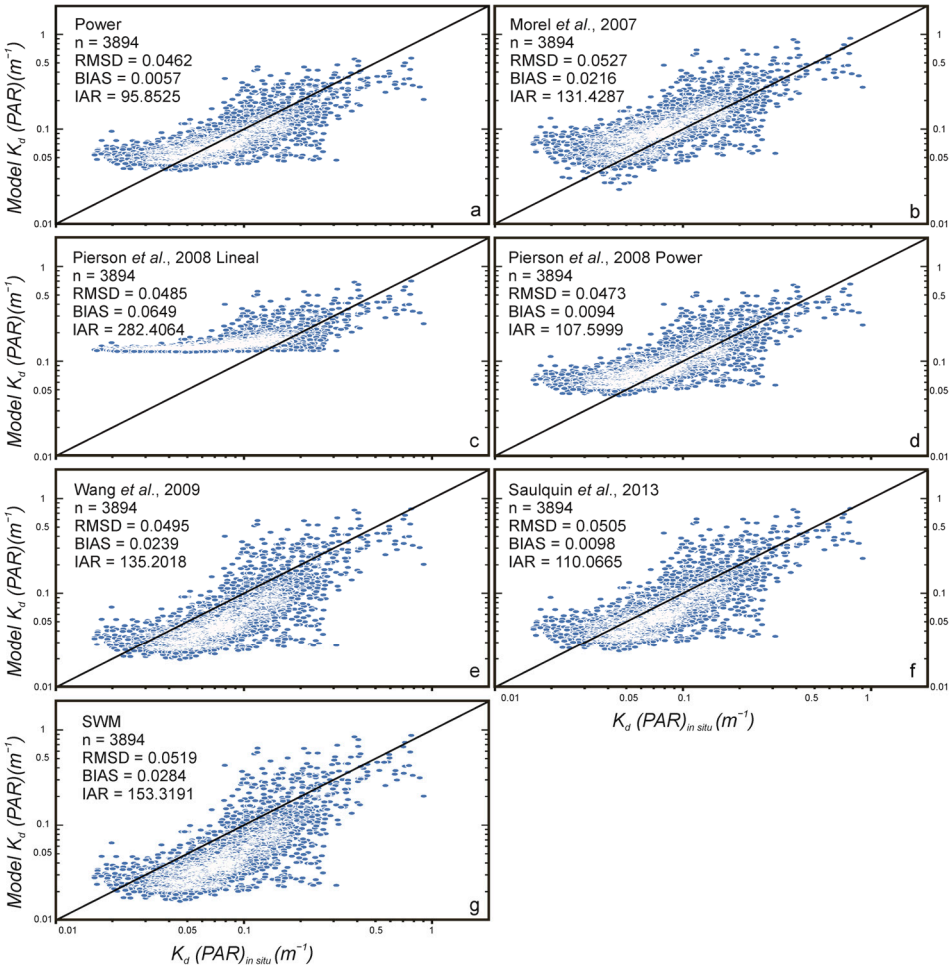


Figure 7. Scatter plots. $K_d(PAR)_{in\ situ}$ and $K_d(PAR)_{model}$. $K_d(490)^{NASA/ESA}_{Standard}$ was used as independent variable in all model. (**a**) Power model (Equation (15)); (**b**) Morel et al. [14] model; (**c**) Pierson et al. [15] model, linear version; (**d**) Pierson et al. [15] model, power version; (**e**) Wang et al., [16]; (**f**) Saulquin et al. [18], (**g**) SWM [37,38] model.

4. Discussion

The reduction of the initial dataset from 18,208 $K_d(PAR)_{in\ situ}$ observations to 11,133 valid measurements demonstrates that acquiring high-quality data on Apparent Optical Properties (AOP) requires strict quality control of environmental conditions during sampling. Factors such as wind speed (due to the formation of white caps), solar position and specific instrument-related considerations are critical to ensuring the accuracy and reliability of the measurements.

The geographical distribution of the 11,133 stations with $K_d(PAR)_{in\ situ}$ (Figure 1) shows global coverage, ensuring that various optical conditions in different marine environments are considered.

The classification of these observations according to the criteria of Castillo-Ramírez et al. [10] (Figure 2) reveals an uneven distribution: Clear Waters ($K_d(PAR) < 0.02$) account for 10,456 observations, Transitional Waters ($K_d(PAR)$ 0.2–0.4) for 534, and Turbid Waters ($K_d(PAR) > 0.4$) for only 143. This latter category, representing approximately 1.3% of

the dataset, confirms the necessity, previously identified by Castillo-Ramírez et al. [10], to enhance in situ sampling efforts in turbid water environments to balance the available observational database.

To ensure the quality of the data used, screening criteria were applied to exclude K_d values below 0.016 m^{-1}, which corresponds to the minimum value recorded for $K_d(490)$ in optically pure seawater [12], as lower values cannot occur under natural conditions. Likewise, values greater than 1 m^{-1} were excluded, following Castillo-Ramírez et al. [10], who established that this value is approximately equivalent to a Z_{SD} of 1.16 m.

Additionally, the exclusion of values greater than 1 m^{-1} was based on previous observations made in the Upper Gulf of California and the Colorado River delta [42–48], all of which reported that satellite data from spectroradiometer sensors are not reliable under high-turbidity conditions. In these regions, when Z_{SD} is less than 1 m, pixels are marked with quality flags, which prevent the generation of ocean color data under these optical conditions.

The relationship between $K_d(PAR)_{in\ situ}$ and $K_d(490)_{in\ situ}$ (Figure 3a) shows a high correlation ($r_{Pearson} = 0.95$, $r_{Spearman} = 0.85$). These results indicate that, following adequate field measurement protocols, $K_d(490)_{in\ situ}$ could be a suitable predictor of $K_d(PAR)_{in\ situ}$. On the other hand, Figure 3b shows the relationship between $K_d(PAR)_{in\ situ}$ and $K_d(490)^{NASA/ESA}_{Standard}$, where a greater dispersion and a lower correlation are observed ($r_{Pearson} = 0.73$; $r_{Spearman} = 0.80$). However, this approach was evaluated based on the criteria of Gregg and Casey [49], Djavidnia et al. [50] and Santamaría-del-Ángel et al. [30], who state that $r_{Pearson}$ values higher than 0.70 indicate a strong association. In this sense, despite the dispersion observed in satellite data, the results suggest that $K_d(490)^{NASA/ESA}_{Standard}$ can be used to estimate $K_d(PAR)$.

The $K_d(PAR)_{in\ situ}$ and $K_d(490)^{NASA/ESA}_{Standard}$ comparisons (Figure 4). comparison allows for identifying potential biases in the relationship between the two variables and exploring possible correction approaches that would reduce the variability of satellite estimates.

The relationship between $K_d(490)_{in\ situ}$ and $K_d(490)^{NASA/ESA}_{Standard}$, (Figure 4) with a correlation values were higher ($r_{Pearson} = 0.78$; $r_{Spearman} = 0.82$) than those observed in Figure 3b. This improvement can be attributed to the fact that in this comparison (Figure 4), both variables represent attenuation at 490 nm, whereas Figure 3b compares $K_d(490)^{NASA/ESA}_{Standard}$ with $K_d(PAR)_{in\ situ}$, a variable that integrates a wider spectral range. Due to this spectral difference, $K_d(PAR)$ is influenced by a greater diversity of optical processes, which introduces greater variability in its relationship with $K_d(490)$. Additionally, the greater dispersion in satellite estimates (Figures 3b and 4) suggests the need to explore nonlinear approaches. This is supported by $r_{Spearman}$, which was higher than $r_{Pearson}$, indicating that the relationship between $K_d(PAR)_{in\ situ}$ and $K_d(490)^{NASA/ESA}_{Standard}$ might not be strictly linear.

It is worth mentioning that in situ values reflect light conditions measured directly in the water column, while satellite estimates represent a broader spatial integration and are subject to other sources of variability. These sources are primarily derived from approximations applied to satellite observations to estimate optical properties of the ocean. Unlike in-situ measurements, where highly accurate point values are obtained, satellite sensors rely on algorithms that interpret the upwelling radiance of water after it has transmitted through the atmosphere.

These algorithms can introduce errors, depending on the coefficients used, the ocean region considered, the variability in water reflectance, the presence of aerosols, and the heterogeneity of the optical properties of the ocean. Although atmospheric corrections could be an additional source of error in satellite products, in practice, most users access

already corrected data. For this reason, a detailed discussion about the improvement of such procedures is beyond the scope of this publication.

The spatial resolution of the satellite sensors also contributes to the observed differences. Although in-situ measurements involve relatively small volumes of water, satellite sensors generate averages of signals over areas of several square kilometers per pixel. This could mask relevant optical gradients, especially in optically complex regions such as the Upper Gulf of California, where it has been documented that optical variability of water can occur on scales smaller than the resolution of satellite remote sensing [44,46–48,51]. Furthermore, the sensor resolution is not the only key factor; the methodology used to obtain the satellite data should also be considered, as discussed below.

Based on the linear relationship illustrated in Figure 3a, one of the first empirical approximations used to estimate $K_d(PAR)$ from $K_d(490)$ was proposed by Kratzer et al. [37], who derived it from 17 observations in the Baltic Sea using a first-order and first-degree model. When evaluating the statistical significance of the model coefficients, these authors found that the intercept was not significant, likely due to the limited number of observations available. That finding led them to propose the Slope-Weighting Model (SWM), which weighs $K_d(490)^{NASA/ESA}_{Standard}$ with the significant slope of the fitted polynomial, under the assumption that $K_d(490)_{in\ situ}$ and $K_d(490)^{NASA/ESA}_{Standard}$ are equivalent. However, as shown in Figure 4 (with axes on a logarithmic scale), this assumption may not be entirely valid.

Soon after, Tang et al. [38] applied the same model in the South China Sea using 30 observations and fitting its own slope. However, unlike Kratzer et al. [37], these authors did not assess the statistical significance of the coefficients and omitted the intercept in their model. These authors used $K_d(490)^{NASA/ESA}_{Standard}$ values obtained from the MODIS-AQUA sensor for their analysis. A key aspect of their study was the warning that the empirical relationships derived from their model would only be valid for the locality in which they were developed, highlighting the importance of validating this type of model under different optical conditions.

The aforementioned analysis, further supported by Figure 3a, suggests the use of linear empirical models to estimate $K_d(PAR)_{in\ situ}$ from $K_d(490)_{in\ situ}$. In Equation (29), the slope coefficient (0.909) was used to develop a version of the global SWM, which could be evaluated and compared with other empirical models reported in the literature (Equation (30)), where $K_d(490)$ represents each of the versions developed in the present study.

To note, in Equation (29), both the intercept and the slope were statistically significant (intercept: $t_{crit} = 1.96$, $t_{calc} = 69.50$; slope: $t_{calc} = 248.59$; $R^2 = 93$; $F_{crit} = 3.84$, $F_{calc} = 61{,}798.74$; $n = 4659$; $\alpha = 5\%$) indicating that the intercept represents the minimum value of $K_d(PAR)$ that cannot be explained by $K_d(490)$. In this sense, considering the $K_d(490)$ reported for optically pure seawater, this minimum value should be approximately 0.016 m^{-1} rather than zero.

It should be noted that in Equation (29), both the intercept and slope were statistically significant (intercept: $t_{crit} = 1.96$, $t_{calc} = 69.50$; slope: $t_{calc} = 248.59$; $R^2 = 93$; $F_{crit} = 3.84$, $F_{calc} = 61{,}798.74$; $n = 4659$; $\alpha = 5\%$), indicating that the intercept represents the minimum value of $K_d(PAR)$ that cannot be explained by $K_d(490)$. In this context, considering the $K_d(490)$ values reported for optically pure seawater, this minimum value should be approximately 0.016 m^{-1} rather than zero.

Previous results (Figures 3 and 4) show that assuming the equivalence of $K_d(490)_{in\ situ}$ with $K_d(490)^{NASA/ESA}_{Standard}$ can introduce variability in the estimation of $K_d(PAR)$ when satellite data are used. Among the sources of variability mentioned above, the coefficients applied to satellite algorithms to derive $K_d(490)$ should be highlighted. Begouen-Demeaux et al. [24] stress the need to review the coefficients used and their temporal

evolution, so they proposed a new corrected version of $K_d(490)^{NASA/ESA}_{Standard}$, which they called $K_d(490)^{NASA/ESA}_{New}$, represented in the BC-ARGOS database.

Although this correction focused on the specific algorithms of each sensor, by observing the linear relationship between $K_d(490)^{NASA/ESA}_{Standard}$ and $K_d(490)^{NASA/ESA}_{New}$ (Figure 5), we explored the possibility of making a similar correction, but using a multisensor approach to improve the consistency of $K_d(490)$ estimates in different satellite missions. This relationship led to the development of Equation (31). The model coefficients were statistically significant, and the relationship showed a nearly perfect fit ($R^2 = 99.9$). This enabled us to obtain a new version of the independent variable named $K_d(490)^{NASA/ESA}_{New}$.

The results obtained (Tables 4–6) show that the coefficients of each model were statistically significant, as were all the approximations evaluated. A key point for these results was the use of a robust database, defined as one that contains a sufficient number of high-quality observations to allow the data to be split into two sets: one for the model construction and the other for independent validation. The greater the number and quality of the data used, the greater the possibility of obtaining statistically significant coefficients in the models. However, by itself, the statistical significance of the coefficients does not guarantee that the model is also significant. Therefore, the global significance of the model must be analyzed by weighing the percentage of the variability of the dependent variable that is explained by that specific model, either using R^2 (Equation (23)) or by the F significance test (Equation (22)).

To select the best model, the significance of its coefficients or the percentage of the variability explained is insufficient. It is essential to carry out a validation through an analysis of residuals, which allows for the evaluation of the actual performance of each model. In this study, validation was carried out based on RMSD (Equation (24)), BIAS (Equation (25)), and MAPE (Equation (26)). Since these descriptors directly depend on the behavior of the residuals, the number of observations used is an important aspect to consider. In this sense, and as mentioned above, having a broad and representative database allows the data to be divided into independent sets and ensures an objective evaluation of the models.

The evaluation of 36 approximations (Figure 6, Appendix B) showed that the power model developed in this study, using $K_d(490)^{NASA/ESA}_{Standard}$ as the independent variable (Equation (15)), returned the best performance. This result suggests that using standard products from different sensors allows for generating longer time series than those produced by a single sensor over its lifetime

The results presented in Figure 7, along with the included residual descriptors ($RMSD$, $BIAS$, and IAR), confirm that the power model (Equation (15)) demonstrated the best performance when compared with the six models reported in the literature using exclusively $K_d(490)^{NASA/ESA}_{Standard}$.

Power models have been employed in different contexts previously. For example, Pierson et al. [15] compared a linear model with a power model in the Baltic Sea and found that the latter yielded a better performance. Likewise, Wang et al. [16] applied a power model for Chesapeake Bay and obtained good fits in turbid waters. Saulquin et al. [18] have also demonstrated that, in environments with $K_d\ (490) > 0.115\ \mathrm{m}^{-1}$, a power model is more suitable for representing the optical variability of water.

In this context, the performance of the best model developed in this study highlights the importance of having a robust database, since a greater number of high-quality observations not only improves the accuracy of the coefficients but also guarantees the representativeness of the model in different optical conditions. The combination of an extensive database and the selection of a power model allowed the development of a

universal model capable of representing the variability of the light field in the ocean under different optical scenarios.

The best model obtained (Equation (15)) will be implemented in the Coastal Marine Information and Analysis System (SIMAR by its Spanish acronym) [52] of the National Commission for the Knowledge and Use of Biodiversity (CONABIO by its Spanish acronym), as part of the commitment to produce knowledge about the short- and long-term changes in Mexican marine-coastal ecosystems using multiscale data time series analysis tools derived from sampling and in-situ monitoring, as well as satellite observations. For the above and in order to identify this product, $K_d(PAR)$ shall be denoted as $K_d(PAR)^{CONABIO/IMAR}$:

$$K_d(PAR)^{CONABIO/IMAR} = 0.575440 \times K_d(490)^{NASA/ESA}_{Standard}{}^{0.683} \quad (34)$$

It is important to note that the six approximations reported in the literature generally returned MPI values below 0.70. This may be because they were developed with databases that do not reflect the global diversity of marine optical conditions. As a result, when they were validated with data involving greater dispersion, they showed lower performance. Additionally, most of these approximations are limited in terms of area and period of application, and work with data from a single sensor. The only exception is the model of Morel et al. [14], who used a database considered robust at the time. However, there are currently almost 40 years of high-quality optical data, which enable a more accurate characterization of marine optical conditions.

The total amount of high-quality data used in this study is the result of a significant sampling effort. Moreover, it is worth noting that the use of optical properties as the basis for marine monitoring systems requires long time series to differentiate natural variability from variability induced by anthropogenic activities [2]. As in-situ monitoring involves considerable logistical and economic effort, remote sensing represents a viable and efficient option. However, to ensure the quality of satellite data, it is necessary to use the latest revisions of standard products, as they include updates to quality flags, which contribute to improving the reliability of the analyses.

The analysis of the in-situ databases and the ocean color missions used in this work (Figure 8) reveals that most databases cannot be maintained sustainably over time. Similarly, satellite missions have a defined lifespan. This discontinuity in both direct measurements and remote observations poses a major challenge for the development of long and consistent time series, which are critical for monitoring the marine environment and detecting long-term trends.

As mentioned above, the use of satellite data for estimating optical properties involves several methodological limitations. As mentioned by Santamaria-del-Ángel et al., [30], comparing a point measurement recorded in the field with an integrated value over a large area captured by satellite remote sensors introduces a source of error that must be considered in the validation of models.

The results obtained in this study confirm that spatial resolution and the methodology for satellite data extraction significantly influence the accuracy of $K_d(PAR)$ estimates. The evolution of ocean color sensors from the Coastal Zone Color Scanner (CZCS) mission to the more recent Plankton, Aerosol, Cloud, ocean Ecosystem (PACE) mission has led to improved resolution of satellite products. Nonetheless, differences in spatial scales remain a critical factor. Most current sensors, such as the Moderate Resolution Imaging Spectroradiometer (MODIS), achieve resolutions of 1 km per pixel (Local Area Cover-LAC); however, the products for large-scale analysis most commonly used correspond to resolutions of 4 km and 9 km per pixel (Global Area Cover-GLO), which impacts the scale of comparison with in-situ data. Furthermore, the availability of images varies according

to the latitude and satellite orbit, which influences the amount of satellite data that matches field measurements, directly affecting the validation of empirical models.

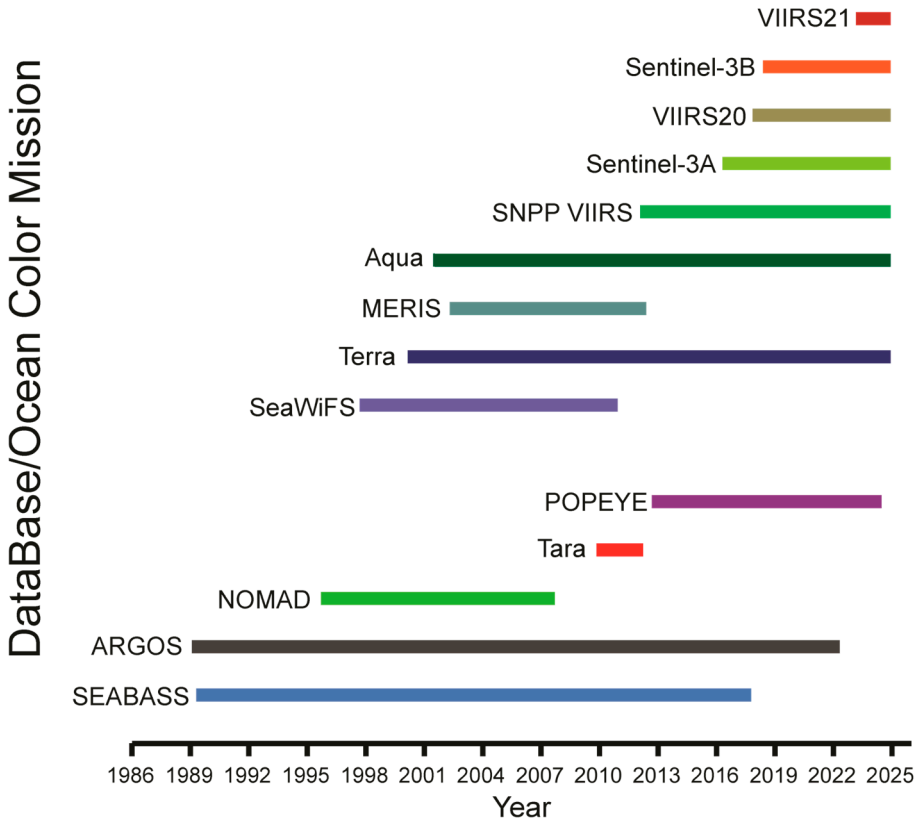


Figure 8. In-situ databases and the ocean color missions used in this work.

Studies such as Tang et al. [38], Wang et al. [16], and Begouen-Demeaux et al. [24] recorded their satellite data in a 5 × 5 pixel box centered on the geographical coordinates of the in-situ station. Although this approach aims to reduce the spatial variability of satellite data, it is essential to note that the area represented by a 5 × 5 pixel box with LAC resolution (1 km pixel size) is 25 km^2. Alternatively, if a 3 × 3 pixel boxes were used, the area represented would be reduced to 9 km^2 [45,47,48,53–56]. This work utilized images with a resolution of 4 km per pixel (GLO); however, instead of applying extraction boxes around the entire area, only the pixel centered on the coordinates of the in-situ station was selected, covering an area of 16 km^2. This approach minimizes spatial overrepresentation and avoids adding values that might skew the relationship between $K_d(490)$ from satellite data and $K_d(PAR)$. In comparison, using 1 km imagery with pixel-based extraction would have involved integrating data from a larger area, which could introduce greater dispersion in model validation. Furthermore, processing 1 km images and focusing on specific coordinates to extract data from a worldwide database would represent a significant challenge in terms of processing time and storage capacity.

Likewise, the use of multisensor approaches proved to be a key tool for improving the spatio-temporal coverage of satellite data and increasing the probability of matches with in situ measurements. Previous studies carried out with the POPEYE database in the Mexican Pacific and the Gulf of California [10,26,44,45,47,48,57,58], found that only 50% of the stations had diurnal in-situ measurements. Of these, only 8% to 15% matched satellite data. This small percentage highlights the importance of integrating data from different sensors to improve the temporal and spatial representativeness of estimates.

Another fundamental aspect in the validation of satellite products is the continuous monitoring of potential deviations in processing algorithms over time. Begouen-Demeaux

et al. [24] emphasize the importance of establishing processes to correct these deviations, which can arise during the sensor's operating time, so that its calibration parameters are modified. Although in this study the corrected versions of $K_d(490)$ from satellite sensors did not show better performance in the models, it is essential to implement cross-recalibration programs between active and inactive sensors. This would improve the quality of satellite products, ensuring their sustained applicability over time.

The results of this study not only have implications for monitoring light attenuation in the ocean but may also contribute to understanding ecological processes for resource management. The variability of the light field in the water column directly influences primary organic productivity [59], which in turn impacts the upper trophic levels [60].

Despite advances in $K_d(PAR)$ estimation based on satellite data, it is essential to recognize that long-term monitoring still depends on the availability and quality of such products. With the recent PACE mission, the accuracy of optical estimates is expected to improve, as this sensor will allow for more detailed spectral integration. However, even if PACE provides a higher spectral resolution, it is essential not to underestimate the importance of historical records captured with previous sensors. With almost four decades of $K_d(490)$ data and other standard products, these records have documented events such as hurricanes [61,62], El Niño [63,64], and the Blob [65,66], providing an invaluable reference for assessing changes in the optical response of the ocean.

Finally, this study highlights the importance of combining various databases and methodological approaches to optimize the estimation of $K_d(PAR)$ using satellite data. Validating empirical models with data from various optical conditions improves the accuracy of estimates and ensures their applicability in different marine environments. The integration of multiple data sources will strengthen marine monitoring systems, providing synoptical tools to assess the marine optical variability in Large Marine Ecosystem. These advances will not only contribute to a better understanding of ocean processes but will also be central in the development of management and conservation strategies for marine ecosystems, facilitating informed and evidence-based decision-making.

5. Conclusions

The results of this study demonstrate that the estimation of $K_d(PAR)$ from $K_d(490)$ derived from satellite sensors can be optimized through empirical models validated with representative data from various optical conditions. Among the approaches evaluated, the power model using $K_d(490)^{NASA/ESA}_{Standard}$ as the independent variable returned the best performance, highlighting the importance of using standard multisensor products to ensure data continuity. The model with the best performance is:

$$K_d(PAR)^{CONABIO/IMAR} = 0.575440 \times K_d(490)^{NASA/ESA\,0.683}_{Standard}$$

Furthermore, this work highlights that methodological differences in the acquisition and resolution of in situ and satellite data can introduce variability in estimates. Factors such as the spatial scale of sensors, the data extraction methodology, and the frequency of observations significantly influence the accuracy of the model. Implementing strategies such as the use of multisensor approaches and cross-calibration between active and inactive sensors is key to improving the reliability of satellite products over time.

The integration of multiple data sources will strengthen marine monitoring systems, providing synoptical tools to assess the marine optical variability in Large Marine Ecosystems. The proposed model will be incorporated into the SIMAR-CONABIO (https://simar.conabio.gob.mx), aiming to facilitate informed decision-making in marine-

coastal ecosystems by communities and governments of 45 countries across the Greater Caribbean and the Northeast Pacific Ocean.

Author Contributions: Conceptualization, E.S.-d.-Á., A.C.-R., S.C.-E., A.G.-S., H.C.-A., J.A.A.-M., R.M.-D., L.R.-d.-l.-C. and M.-T.S.-F.; methodology, J.A.Z.-H., E.S.-d.-Á., A.C.-R., S.C.-E., A.G.-S., H.C.-A., J.A.A.-M., R.M.-D., L.R.-d.-l.-C. and M.-T.S.-F.; software, E.S.-d.-Á. and A.C.-R.; validation, E.S.-d.-Á. and A.C.-R.; formal analysis, J.A.Z.-H., E.S.-d.-Á., A.C.-R., S.C.-E., A.G.-S., H.C.-A., J.A.A.-M., R.M.-D., L.R.-d.-l.-C. and M.-T.S.-F.; investigation, J.A.Z.-H., E.S.-d.-Á., A.C.-R., S.C.-E., A.G.-S., H.C.-A., J.A.A.-M., R.M.-D., L.R.-d.-l.-C. and M.-T.S.-F.; resources, E.S.-d.-Á. and A.C.-R.; data curation, J.A.Z.-H., E.S.-d.-Á. and A.C.-R.; writing—original draft preparation, J.A.Z.-H., E.S.-d.-Á., A.C.-R. and M.-T.S.-F.; writing—review and editing, E.S.-d.-Á., A.C.-R., S.C.-E., A.G.-S., H.C.-A., J.A.A.-M., R.M.-D., L.R.-d.-l.-C. and M.-T.S.-F.; visualization, E.S.-d.-Á., A.C.-R. and M.-T.S.-F.; supervision, E.S.-d.-Á., A.C.-R., S.C.-E. and M.-T.S.-F.; project administration, S.C.-E., E.S.-d.-Á., H.C.-A., R.M.-D. and L.R.-d.-l.-C.; funding acquisition, S.C.-E. All authors have read and agreed to the published version of the manuscript.

Funding: This research was conducted within the framework of the project "Implementation of the Strategic Action Program (SAP) for the Gulf of Mexico Large Marine Ecosystem (GoM-LME)" (ID 6952), funded by the Global Environment Facility (GEF). The project is implemented by the United Nations Environment Programme (UNEP) and executed by the United Nations Office for Project Services (UNOPS). It received in-kind co-financing from the National Oceanic and Atmospheric Administration (NOAA) and the Ministry of Environment and Natural Resources of Mexico (SEMARNAT), as well as from the Cartagena Convention and its Caribbean Regional Coordination Unit (CAR/RCU).

Data Availability Statement: Publicly available datasets were analyzed in this study. These data are available from the following sources: Sea-BASS [21], NOMAD version 2.0 [22], the Tara Oceans Consortium dataset [23], and the BGC-Argo dataset (version 2) [24]. The Phytoplankton Ecology Team (POPEYE) dataset is available from the Autonomous University of Baja California (UABC) upon request.

Acknowledgments: The authors thank the National Trust Fund for Biodiversity at CONABIO for their administrative support. We also thank the Secretariat of Science, Humanities, Technology and Innovation (SECIHTI) for funding under project CF-2019-1327711, and the NASA Ocean Biology Processing Group (OBPG) for providing the satellite data. The authors appreciate the assistance of María Elena Sánchez-Salazar with the English translation of this manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AOP	Apparent Optical Properties
BGC-Argo	Biogeochemical-Argo
BIAS	Bias Analysis
b_k	Coefficient associated with the variable independent k
CAR/RCU	Cartagena Convention and the Regional Coordination Unit
CONABIO	National Commission for the Knowledge and Use of Biodiversity
$Cov_{a,b}$	Covariance of the a and b
$Cov_{Ra,Rb}$	Covariance of the ranks of a and b
CZCS	Coastal Zone Color Scanner
E_{0-}	Downwelling irradiance just below the sea surface
$E_d(PAR)$	Downwelling PAR irradiance
$E_d(\lambda)$	Downwelling spectral irradiance
ES	Ecosystem services
F_{cal}	Calculated F-test value to test the overall significance of each model

GEF	Global Environment Facility
GLO	Global Area Coverage
IAR	Integrated Absolute Residuals
IV	Independent variable
K_d	Diffuse light attenuation coefficient
K_{dw}	Diffuse light attenuation coefficient of optically pure seawater
$K_d\ (PAR)$	Diffuse attenuation coefficient of photosynthetically active radiation
$K_d(PAR)_{in\ situ}$	Diffuse attenuation coefficient of photosynthetically available radiation in situ data
$K_d(PAR)_{model}$	Diffuse attenuation coefficient of photosynthetically active radiation modeling
$K_d\ (490)$	Diffuse attenuation coefficient at 490 nm coefficient
$K_d(490)_{in\ situ}$	Diffuse attenuation coefficient at 490 nm coefficient in situ data
$K_d(490)^{NASA/ESA}_{Standard}$	$K_d\ (490)$ satellite standard data
$K_d(490)^{NASA/ESA}_{New}$	recalculation of $K_d(490)^{NASA/ESA}_{Standard}$ according with Begouen-Demeaux et al. [24]
L2	Level 2
L3	Level 3
LAC	Local Area Coverage
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MERIS	Medium Resolution Imaging Spectrometer
MLAC	Merged Local Area Coverage
MODIS	Moderate-Resolution Imaging Spectroradiometer
MPI	Model Performance Index
NESA	NASA-ESA Inspired Ocean Color Model
NOAA	National Oceanic and Atmospheric Administration
NOMAD	NASA bio-Optical Marine Algorithm Dataset
OLCI	Ocean and Land Colour Instrument
PACE	Plankton, Aerosol, Cloud, ocean Ecosystem
PAR	Photosynthetically Active Radiation
POPEYE	Phytoplankton Ecology Team of the UABC
Power	Power Regression Model
R^2	Coefficient of determination
$R_{\|BIAS\|}$	Individual absolute ranks of $BIAS$
R_{MAPE}	Individual ranks for $MAPE$
RMSD	Root Mean Square Deviation
R_{RMSD}	Individual ranks for $RMSD$
$r_{Pearson}$	Pearson's correlation coefficient
$r_{Spearman}$	Spearman's correlation coefficient
SeaBASS	SeaWiFS Bio-optical Archive and Storage System
SeaWiFS	Sea-viewing Wide Field-of-view Sensor
SECIHTI	Secretary of Science, Humanities, Technology and Innovation
SE_{bk}	Standard error of the coefficient b_k
SDG	Sustainable Development Goals
SEMARNAT	Secretary of Environment and Natural Resources
SIMAR	Coastal Marine Information and Analysis System
$StDev_j$	Standard deviations of j
$StDev_{Rj}$	Standard deviation of the ranks of j
SWM	Slope Weighting Model
UABC	Autonomous University of Baja California
UNEP	United Nations Environmental Programme
VIIRS	Visible Infrared Imaging Radiometer Suite
Z_{SD}	Secchi Disk depth

Appendix A. Global Distribution of Stations with $K_d(PAR)_{insitu}$ Data

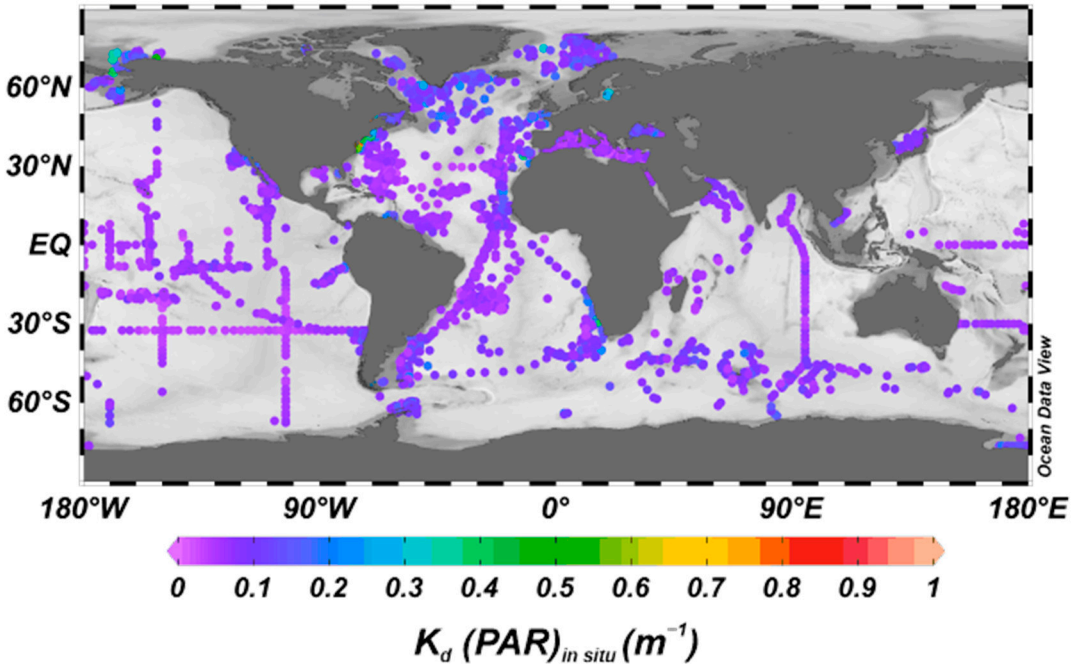
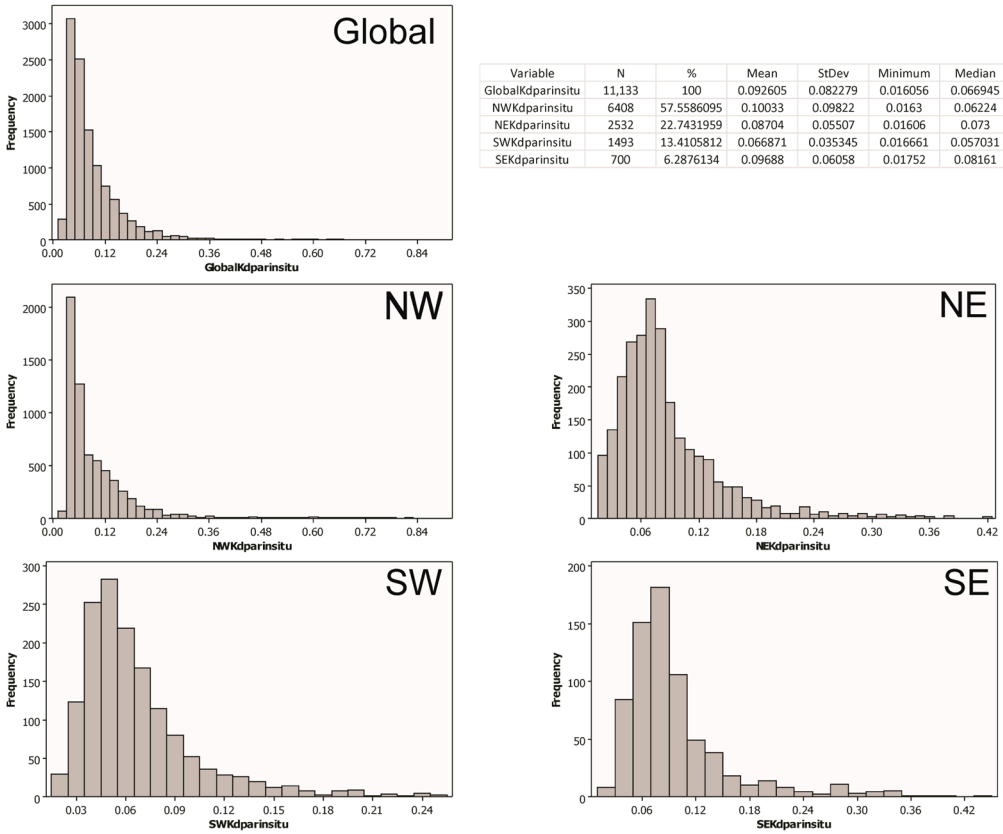


Figure A1. Global distribution of stations with $K_d(PAR)_{in\ situ}$ data. The color bar indicates the variability of $K_d(PAR)$.



Variable	N	%	Mean	StDev	Minimum	Median	Maximum
GlobalKdparinsitu	11,133	100	0.092605	0.082279	0.016056	0.066945	0.9022
NWKdparinsitu	6408	57.5586095	0.10033	0.09822	0.0163	0.06224	0.9022
NEKdparinsitu	2532	22.7431959	0.08704	0.05507	0.01606	0.073	0.41713
SWKdparinsitu	1493	13.4105812	0.066871	0.035345	0.016661	0.057031	0.2528
SEKdparinsitu	700	6.2876134	0.09688	0.06058	0.01752	0.08161	0.43838

Figure A2. Histograms of the $K_d(PAR)_{in\ situ}$ the database used. The total data histogram (Global) is shown, as well as histograms by hemisphere (NW, NE, SW, and SE). A table of the data descriptors used in each histogram is also included.

Appendix B

Table A1. Performance of the linear, power, and NESA empirical models developed in this work and those reported in the literature for the estimation of $K_d(PAR)$ from different versions of K_d (490). IV, independent variable.

Model	IV	RMSD	BIAS	MAPE	MPI
Linear	$K_d(490)^{NASA/ESA}_{Standard}$	0.0485	0.0018	32.329	0.7593
Power	$K_d(490)^{NASA/ESA}_{Standard}$	0.0462	0.0057	26.856	0.8704
NESA	$K_d(490)^{NASA/ESA}_{Standard}$	0.0483	0.0083	25.825	0.8333
Morel et al. [14]	$K_d(490)^{NASA/ESA}_{Standard}$	0.0527	0.0216	44.487	0.2407
Pierson et al. [15]	$K_d(490)^{NASA/ESA}_{Standard}$	0.0485	0.0649	126.949	0.2963
Pierson et al. [15]	$K_d(490)^{NASA/ESA}_{Standard}$	0.0473	0.0094	35.672	0.6759
Wang et al. [16]	$K_d(490)^{NASA/ESA}_{Standard}$	0.0495	0.0239	36.811	0.4537
Saulquin et al. [18]	$K_d(490)^{NASA/ESA}_{Standart}$	0.0505	0.0098	29.551	0.5648
SWM [37,38]	$K_d(490)^{NASA/ESA}_{Standard}$	0.0519	0.0284	43.185	0.2685
Linear	$K_d(490)^{NASA/ESA}_{New}$	0.0529	0.0002	31.853	0.5463
Power	$K_d(490)^{NASA/ESA}_{New}$	0.0494	0.0042	27.032	0.7593
NESA	$K_d(490)^{NASA/ESA}_{New}$	0.0506	0.0045	26.991	0.6574
Morel et al. [14]	$K_d(490)^{NASA/ESA}_{New}$	0.0517	0.0098	34.766	0.4630
Pierson et al. [15]	$K_d(490)^{NASA/ESA}_{New}$	0.0524	0.0593	122.131	0.1019
Pierson et al. [15]	$K_d(490)^{NASA/ESA}_{New}$	0.0487	0.0004	30.166	0.8056
Wang et al. [16]	$K_d(490)^{NASA/ESA}_{New}$	0.0506	0.0320	41.253	0.3148
Saulquin et al. [18]	$K_d(490)^{NASA/ESA}_{New}$	0.0504	0.0192	31.726	0.4907
SWM [37,38]	$K_d(490)^{NASA/ESA}_{New}$	0.0520	0.0366	47.946	0.1667
Linear	$K_d(490)\,_{Corr}^{NASA-ESA/Standard}$	0.0504	0.0020	30.680	0.6667
Power	$K_d(490)\,_{Corr}^{NASA-ESA/Standard}$	0.0600	0.0405	67.424	0.1019
NESA	$K_d(490)\,_{Corr}^{NASA-ESA/Standard}$	0.0488	0.0077	26.199	0.7778
Morel et al. [14]	$K_d(490)\,_{Corr}^{NASA-ESA/Standard}$	0.0488	0.0179	45.979	0.4630
Pierson et al. [15]	$K_d(490)\,_{Corr}^{NASA-ESA/Standard}$	0.0525	0.0612	126.953	0.0648
Pierson et al. [15]	$K_d(490)\,_{Corr}^{NASA-ESA/Standard}$	0.0482	0.0056	36.846	0.7037
Wang et al. [16]	$K_d(490)\,_{Corr}^{NASA-ESA/Standard}$	0.0492	0.0288	35.571	0.4815
Saulquin et al. [18]	$K_d(490)\,_{Corr}^{NASA-ESA/Standard}$	0.0487	0.0147	27.793	0.6944
SWM [37,38]	$K_d(490)\,_{Corr}^{NASA-ESA/Standard}$	0.0497	0.0339	41.817	0.3426
Linear	$K_d(490)\,_{Corr}^{NASA-ESA/New}$	0.0512	0.0016	31.434	0.6204
Power	$K_d(490)\,_{Corr}^{NASA-ESA/New}$	0.0682	0.0555	84.859	0.0741
NESA	$K_d(490)\,_{Corr}^{NASA-ESA/New}$	0.0488	0.0062	26.544	0.7685
Morel et al. [14]	$K_d(490)\,_{Corr}^{NASA-ESA/New}$	0.0499	0.0190	46.377	0.3704
Pierson et al. [15]	$K_d(490)\,_{Corr}^{NASA-ESA/New}$	0.0519	0.0620	127.223	0.0741
Pierson et al. [15]	$K_d(490)\,_{Corr}^{NASA-ESA/New}$	0.0483	0.0065	37.069	0.6481
Wang et al. [16]	$K_d(490)\,_{Corr}^{NASA-ESA/New}$	0.0497	0.0277	35.651	0.4537
Saulquin et al. [18]	$K_d(490)\,_{Corr}^{NASA-ESA/New}$	0.0495	0.0135	27.952	0.6296
SWM [37,38]	$K_d(490)\,_{Corr}^{NASA-ESA/New}$	0.0506	0.0327	41.932	0.2963

References

1. UN (United Nations). Available online: https://www.un.org/sustainabledevelopment/es/development-agenda/ (accessed on 2 October 2025).
2. Santamaría-del-Ángel, E.; Sebastia-Frasquet, M.T.; Millán-Nuñez, R.; González-Silvera, A.; Cajal-Medrano, R. Anthropocentric BIAS in Management Policies. Are We Efficiently Monitoring Our Ecosystems. In *Coastal Ecosystems: Experiences and Recommenda-*

tions for Environmental Monitoring Programs; Sebastia-Frasquet, M.T., Ed.; Nova Science Publishers: New York, NY, USA, 2015; pp. 1–12.
3. Häder, D.-P.; Banaszak, A.T.; Villafañe, V.E.; Narvarte, M.A.; González, R.A.; Helbling, E.W. Anthropogenic Pollution of Aquatic Ecosystems: Emerging Problems with Global Implications. *Sci. Total Environ.* **2020**, *713*, 136586. [CrossRef]
4. Mojiri, A.; Zhou, J.L.; Robinson, B.; Ohashi, A.; Ozaki, N.; Kindaichi, T.; Farraji, H.; Vakili, M. Pesticides in Aquatic Environments and Their Removal by Adsorption Methods. *Chemosphere* **2020**, *253*, 126646. [CrossRef]
5. Santamaría-del-Ángel, E.; Soto, I.; Millán-Núñez, R.; González, A.; Wolny, J.; Cerdeira-Estrada, S.; Cajal-Medrano, R.; Muller, F.; Padilla-Rosas, Y.X.S.; Mercado-Santana, A. Phytoplankton Blooms: New Initiative Using Marine Optics as a Basis for Monitoring Programs. In *Coastal Ecosystems: Experiences and Recommendations for Environmental Monitoring Programs*; Sebastia-Frasquet, M.T., Ed.; Nova Science Publishers: New York, NY, USA, 2015; pp. 57–88.
6. Kirk, J.T.O. *Light and Photosynthesis in Aquatic Ecosystems*, 3rd ed.; Cambridge University Press: Cambridge, UK, 2011; p. 649. ISBN 978-1-139-16821-2.
7. Lee, Z.; Du, K.; Arnone, R.; Liew, S.; Penta, B. Penetration of Solar Radiation in the Upper Ocean: A Numerical Model for Oceanic and Coastal Waters. *J. Geophys. Res. Oceans* **2005**, *110*, 1–12. [CrossRef]
8. Lin, J.; Lee, Z.; Ondrusek, M.; Kahru, M. Attenuation Coefficient of Usable Solar Radiation of the Global Oceans. *J. Geophys. Res. Oceans* **2016**, *121*, 3228–3236. [CrossRef]
9. Harvey, E.T.; Walve, J.; Andersson, A.; Karlson, B.; Kratzer, S. The Effect of Optical Properties on Secchi Depth and Implications for Eutrophication Management. *Front. Mar. Sci.* **2019**, *5*, 496. [CrossRef]
10. Castillo-Ramírez, A.; Santamaría-del-Ángel, E.; González-Silvera, A.; Frouin, R.; Sebastiá-Frasquet, M.T.; Tan, J.; Lopez-Calderon, J.; Sánchez-Velasco, L.; Enríquez-Paredes, L. A New Algorithm to Estimate Diffuse Attenuation Coefficient from Secchi Disk Depth. *J. Mar. Sci. Eng.* **2020**, *8*, 558. [CrossRef]
11. IOCCG. *Uncertainties in Ocean Colour Remote Sensing*; Mélin, F., Ed.; IOCCG Report Series, No. 18; International Ocean Colour Coordinating Group (IOCCG): Dartmouth, NS, Canada, 2019; ISBN 978-1-896246-68-0.
12. Mueller, J.L. SeaWiFS Algorithm for the Diffuse Attenuation Coefficient, K(490), Using Water-Leaving Radiances at 490 and 555 nm. In *SeaWiFS Postlaunch Calibration and Validation Analyses*; Standford, B.H., Ed.; NASA Goddard Space Flight Center: Washington, DC, USA, 2000; Volume 3, pp. 24–27.
13. Mobley, C.D. *Light and Water: Radiative Transfer in Natural Waters*; Academic Press: New York, NY, USA, 1994; p. 592.
14. Morel, A.; Huot, Y.; Gentili, B.; Werdell, P.J.; Hooker, S.B.; Franz, B.A. 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.* **2007**, *111*, 69–88. [CrossRef]
15. Pierson, D.C.; Kratzer, S.; Strömbeck, N.; Håkansson, B. Relationship between the Attenuation of Downwelling Irradiance at 490 Nm with the Attenuation of PAR (400 nm–700 nm) in the Baltic Sea. *Remote Sens. Environ.* **2008**, *112*, 668–680. [CrossRef]
16. Wang, M.; Son, S.; Harding, L.W., Jr. Retrieval of Diffuse Attenuation Coefficient in the Chesapeake Bay and Turbid Ocean Regions for Satellite Ocean Color Applications. *J. Geophys. Res. Oceans* **2009**, *114*, 1–15. [CrossRef]
17. Shang, S.; Lee, Z.; Wei, G. Characterization of MODIS-Derived Euphotic Zone Depth: Results for the China Sea. *Remote Sens. Environ.* **2011**, *115*, 180–186. [CrossRef]
18. Saulquin, B.; Hamdi, A.; Gohin, F.; Populus, J.; Mangin, A.; d'Andon, O.F. Estimation of the Diffuse Attenuation Coefficient KdPAR Using MERIS and Application to Seabed Habitat Mapping. *Remote Sens. Environ.* **2013**, *128*, 224–233. [CrossRef]
19. Jerlov, N.G. *Marine Optics*, 2nd ed.; American Elsevier Publishing Company Incorporation: New York, NY, USA, 1976; Volume 14, ISBN 978-0-08-087050-2.
20. Solonenko, M.G.; Mobley, C.D. Inherent Optical Properties of Jerlov Water Types. *Appl. Opt.* **2015**, *54*, 5392–5401. [CrossRef] [PubMed]
21. Werdell, P.J.; Bailey, S.; Fargion, G.; Pietras, C.; Knobelspiesse, K.; Feldman, G.; McClain, C. Unique Data Repository Facilitates Ocean Color Satellite Validation. *Eos. Trans. AGU* **2003**, *84*, 377–387. [CrossRef]
22. Werdell, P.J.; Bailey, S.W. An Improved In-Situ Bio-Optical Data Set for Ocean Color Algorithm Development and Satellite Data Product Validation. *Remote Sens. Environ.* **2005**, *98*, 122–140. [CrossRef]
23. Tara Oceans Consortium, Coordinators; Tara Oceans Expedition, Participants. Environmental Context of All Samples from the Tara Oceans Expedition (2009–2013), about Mesoscale Features [Dataset]. *PANGAEA* **2016**. [CrossRef]
24. Begouen-Demeaux, C.; Boss, E.; Tan, J.; Frouin, R. Algorithms to Retrieve the Spectral Diffuse Attenuation Coefficient of Light in the Ocean from Remote Sensing. *Opt. Express* **2024**, *32*, 2507–2526. [CrossRef]
25. Preisendorfer, R.W. Secchi Disk Science: Visual Optics of Natural Waters. *Limnol. Oceanogr.* **1986**, *31*, 909–926. [CrossRef]
26. Castillo-Ramírez, A.; Santamaría-del-Ángel, E.; González-Silvera, A.; Aguilar-Maldonado, J.; Lopez-Calderon, J.; Sebastiá-Frasquet, M.-T. Use of Digital Images as a Low-Cost System to Estimate Surface Optical Parameters in the Ocean. *Sensors* **2023**, *23*, 3199. [CrossRef]
27. NASA Ocean Color Web. Available online: https://oceancolor.gsfc.nasa.gov/ (accessed on 2 October 2025).

28. Kahru, M.; Di Lorenzo, E.; Manzano-Sarabia, M.; Mitchell, B.G. Spatial and Temporal Statistics of Sea Surface Temperature and Chlorophyll Fronts in the California Current. *J. Plankton Res.* **2012**, *34*, 749–760. [CrossRef]
29. Kahru, M.; Kudela, R.M.; Anderson, C.R.; Mitchell, B.G. Optimized Merger of Ocean Chlorophyll Algorithms of MODIS-Aqua and VIIRS. *GRSL* **2015**, *12*, 2282–2285. [CrossRef]
30. Santamaría-del-Ángel, E.; Millan-Nuñez, R.; Gonzalez-Silvera, A.; Cajal, R. Comparison of In Situ and Remotely-Sensed Chl-a Concentrations: A Statistical Examination of the Match-up Approach. In *Handbook of Satellite Remote Sensing Image Interpretation: Applications for Marine Living Resources Conservation and Management*; Morales, J., Stuart, V., Platt, T., Sathyendranath, S., Eds.; EU PRESPO and IOCCG: Dartmouth, NS, Canada, 2011; pp. 241–260.
31. Perdices, M. Null Hypothesis Significance Testing, p-Values, Effects Sizes and Confidence Intervals. *Brain Impair.* **2018**, *19*, 70–80. [CrossRef]
32. Mu, Y.; Liu, X.; Wang, L. A Pearson's Correlation Coefficient Based Decision Tree and Its Parallel Implementation. *Inf. Sci.* **2018**, *435*, 40–58. [CrossRef]
33. Farnsworth, D.L.; Triola, M.F. Review of Elementary Statistics. *Technometrics* **1990**, *32*, 456–457. [CrossRef]
34. Kokoska, S.; Nevison, C. Critical Values For Spearman's Rank Correlation Coefficient. In *Statistical Tables and Formulae*; Kokoska, S., Nevison, C., Eds.; Springer: New York, NY, USA, 1989; p. 86. ISBN 978-1-4613-9629-1.
35. IOCCG. *Synergy between Ocean Colour and Biogeochemical/Ecosystem Models*; Dutkiewicz, S., Ed.; IOCCG Report Series, No. 19; International Ocean Colour Coordinating Group (IOCCG): Dartmouth, NS, Canada, 2020; p. 184.
36. Luijken, K.; Wynants, L.; van Smeden, M.; Calster, B.V.; Steyerberg, E.W.; Groenwold, R.H.H.; Timmerman, D.; Bourne, T.; Ukaegbu, C. Changing Predictor Measurement Procedures Affected the Performance of Prediction Models in Clinical Examples. *J. Clin. Epidemiol.* **2020**, *119*, 7–18. [CrossRef] [PubMed]
37. Kratzer, S.; Håkansson, B.; Sahlin, C. Assessing Secchi and Photic Zone Depth in the Baltic Sea from Satellite Data. *Ambio* **2003**, *32*, 577–585. [CrossRef]
38. Tang, S.; Chen, C.; Zhan, H.; Xu, D. Remotely-Sensed Estimation of the Euphotic Depth in the Northern South China Sea. In Proceedings of the 2007 IEEE International Geoscience and Remote Sensing Symposium, Barcelona, Spain, 23–28 July 2007; pp. 917–920.
39. Araveeporn, A. Improved Probability-Weighted Moments and Two-Stage Order Statistics Methods of Generalized Extreme Value Distribution. *Mathematics* **2025**, *13*, 2295. [CrossRef]
40. Jui, J.J.; Molla, M.M.I.; Ahmad, M.A.; Hettiarachchi, I.T. Recent Advances and Applications of the Multi-Verse Optimiser Algorithm: A Survey from 2020 to 2024. *Arch. Comput. Methods Eng.* **2025**, *32*, 4491–4524. [CrossRef]
41. Wilcoxon, F. Individual Comparisons by Ranking Methods. *Biom. Bull.* **1945**, *1*, 80–83. [CrossRef]
42. Santamaría-del-Ángel, E.; Millán-Núñez, R.; De-Peña-Nettel, G. Effect of Turbidity on Primary Productivity at Two Stations in the Area of the Colorado River Delta. *Cienc. Mar.* **1996**, *22*, 483–493. [CrossRef]
43. Millán-Nuñez, R.; Santamaría-del-Ángel, E.; Cajal-Medrano, R.; Barocio-León, O.A. The Colorado River Delta: A High Primary Productivity Ecosystem. *Cienc. Mar.* **1999**, *25*, 509–524. [CrossRef]
44. Bastidas-Salamanca, M.; Gonzalez-Silvera, A.; Millán-Núñez, R.; Santamaria-del-Angel, E.; Frouin, R. Bio-Optical Characteristics of the Northern Gulf of California during June 2008. *Int. J. Oceanogr.* **2014**, *2014*, 384618. [CrossRef]
45. Mercado-Santana, J.A.; Santamaría-del-Ángel, E.; González-Silvera, A.; Sánchez-Velasco, L.; Gracia-Escobar, M.F.; Millán-Núñez, R.; Torres-Navarrete, C. Productivity in the Gulf of California Large Marine Ecosystem. *Environ. Dev.* **2017**, *22*, 18–29. [CrossRef]
46. Aguilar-Maldonado, J.A.; Santamaría-del-Ángel, E.; Sebastiá-Frasquet, M.T. Reflectances of SPOT Multispectral Images Associated with the Turbidity of the Upper Gulf of California. *Rev. De Teledetección* **2017**, *50*, 1–16. [CrossRef]
47. Betancur-Turizo, S.P.; González-Silvera, A.G.; Santamaría-Del-Ángel, E.; Millán-Núñez, R.; Millán-Núñez, E.; García-Nava, H.; Godínez, V.M.; Sánchez-Velasco, L. Variability in the Light Absorption Coefficient by Phytoplankton, Non-Algal Particles and Colored Dissolved Organic Matter in the Northern Gulf of California. *Open J. Mar. Sci.* **2018**, *8*, 20–37. [CrossRef]
48. Betancur-Turizo, S.P.; González-Silvera, A.; Santamaría-del-Ángel, E.; Tan, J.; Frouin, R. Evaluation of Semi-Analytical Algorithms to Retrieve Particulate and Dissolved Absorption Coefficients in Gulf of California Optically Complex Waters. *Remote Sens.* **2018**, *10*, 1443. [CrossRef]
49. Gregg, W.W.; Casey, N.W. Global and Regional Evaluation of the SeaWiFS Chlorophyll Data Set. *Remote Sens. Environ.* **2004**, *93*, 463–479. [CrossRef]
50. Djavidnia, S.; Mélin, F.; Hoepffner, N. Analysis of Multi-Sensor Global and Regional Ocean Colour Products. *MERSEA-IP Mar. Environ. Secur. Eur. Area-Integr. Proj. Rep. Deliv. D* **2006**, *2*, 228.
51. Aguilar-Maldonado, J.A.; Santamaría-del-Ángel, E.; Sebastiá-Frasquet, M.T. Applying SPOT Images to Study the Colorado River Effects on the Upper Gulf of California. *Proceedings* **2017**, *2*, 182. [CrossRef]
52. SIMAR. Available online: https://simar.conabio.gob.mx/explorer/ (accessed on 2 October 2025).

53. Aguilar-Maldonado, J.A.; Santamaría-del-Ángel, E.; González-Silvera, A.; Cervantes-Rosas, O.D.; Sebastiá-Frasquet, M.-T. Mapping Satellite Inherent Optical Properties Index in Coastal Waters of the Yucatán Peninsula (Mexico). *Sustainability* **2018**, *10*, 1894. [CrossRef]
54. Aguilar-Maldonado, J.A.; Santamaría-del-Ángel, E.; González-Silvera, A.; Cervantes-Rosas, O.D.; López, L.M.; Gutiérrez-Magness, A.; Cerdeira-Estrada, S.; Sebastiá-Frasquet, M.-T. Identification of Phytoplankton Blooms under the Index of Inherent Optical Properties (IOP Index) in Optically Complex Waters. *Water* **2018**, *10*, 129. [CrossRef]
55. Santamaría-del-Ángel, E.; Cañon-Páez, M.L.; Sebastiá-Frasquet, M.T.; González-Silvera, A.; Gutierrez, A.L.; Aguilar-Maldonado, J.A.; López-Calderón, J.; Camacho-Ibar, V.; Franco-Herrera, A.; Castillo-Ramírez, A. Interannual Climate Variability in the West Antarctic Peninsula under Austral Summer Conditions. *Remote Sens.* **2021**, *13*, 1122. [CrossRef]
56. Mirabal, U.; Linacre, L.; Durazo, R.; Santamaría-del-Ángel, E.; Pallàs-Sanz, E.; Lara-Lara, J.R. Regionalization of Oceanic Waters Based on Satellite Bio-Optical Properties in the Central and Southern Gulf of Mexico. *RSASE* **2025**, *39*, 101695. [CrossRef]
57. Miranda-Alvarez, C.; González-Silvera, A.; Santamaría-del-Angel, E.; López-Calderón, J.; Godínez, V.M.; Sánchez-Velasco, L.; Hernández-Walls, R. Phytoplankton Pigments and Community Structure in the Northeastern Tropical Pacific Using HPLC-CHEMTAX Analysis. *J. Oceanogr.* **2020**, *76*, 91–108. [CrossRef]
58. Larios-Muñiz, M.; González-Silvera, A.; Castro, R.; Santamaría-del-Ángel, E.; Collins, C.A.; López-Calderón, J. Variability of Hydrographic Factors, Biomass and Structure of the Phytoplankton Community at the Entrance to the Gulf of California (Spring 2013). *Cont. Shelf Res.* **2022**, *235*, 104665. [CrossRef]
59. Rautenberger, R.; Fernández, P.A.; Strittmatter, M.; Heesch, S.; Cornwall, C.E.; Hurd, C.L.; Roleda, M.Y. Saturating Light and Not Increased Carbon Dioxide under Ocean Acidification Drives Photosynthesis and Growth in Ulva Rigida (Chlorophyta). *Ecol. Evol.* **2015**, *5*, 874–888. [CrossRef]
60. Ruiz-Cooley, R.I.; Gerrodette, T.; Fiedler, P.C.; Chivers, S.J.; Danil, K.; Ballance, L.T. Temporal Variation in Pelagic Food Chain Length in Response to Environmental Change. *Sci. Adv.* **2017**, *3*, e1701140. [CrossRef]
61. Liu, Y.; He, R.; Lee, Z. Effects of Ocean Optical Properties and Solar Attenuation on the Northwestern Atlantic Ocean Heat Content and Hurricane Intensity. *Geophys. Res. Lett.* **2021**, *48*, e2021GL094171. [CrossRef]
62. Shi, W.; Wang, M. Ocean Responses to Hurricane Ian from Daily Gap-Free Satellite Observations. *Remote Sens. Lett.* **2023**, *14*, 877–889. [CrossRef]
63. Palmer, P.I. The Role of Satellite Observations in Understanding the Impact of El Niño on the Carbon Cycle: Current Capabilities and Future Opportunities. *Philos. Trans. R. Soc. B Biol. Sci.* **2018**, *373*, 20170407. [CrossRef] [PubMed]
64. Jauregui, Y.R.; Chen, S.S. MJO-Induced Warm Pool Eastward Extension Prior to the Onset of El Niño: Observations from 1998 to 2019. *J. Clim.* **2024**, *37*, 855–873. [CrossRef]
65. Shen, X.; Zhan, W.; Zhang, Y.; He, Q.; Bo, Y.; Liu, Y.; Zhan, H. Spatial Heterogeneity and Seasonality of Phytoplankton Responses to Marine Heatwaves in the Northeast Pacific. *Environ. Res. Lett.* **2024**, *20*, 014042. [CrossRef]
66. Chen, H.-H.; Wang, Y.; Li, X.; Wan, L.; Yuan, Y.; Yan, Y.; Hannah, C.; Chai, F. Arctic Warming as a Potential Trigger for the Warm Blob in the Northeast Pacific. *npj Clim. Atmos. Sci.* **2025**, *8*, 111. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.