Scientific Research Publishing

Advances in Remote Sensing, 2025, 14(3), 147-163
https://www.scirp.org/journal/ars
ISSN Online: 2169-2688
ISSN Print: 2169-267X

Optical Classification of Water Types in Cook Inlet, Alaska

Hisatomo Waga[1,2*], Mark A. Johnson[3]

[1]International Arctic Research Center, University of Alaska Fairbanks, Alaska, USA
[2]International Polar and Earth Environmental Research Center, National Institute of Polar Research, Tokyo, Japan
[3]College of Fisheries and Ocean Sciences, University of Alaska Fairbanks, Alaska, USA
Email: *hwaga@alaska.edu

How to cite this paper: Waga, H. and Johnson, M.A. (2025) Optical Classification of Water Types in Cook Inlet, Alaska. *Advances in Remote Sensing*, **14**, 147-163.
https://doi.org/10.4236/ars.2025.143009

Received: July 29, 2025
Accepted: September 16, 2025
Published: September 19, 2025

Abstract

Ocean color is determined by the complex interactions of incident light with the optical properties of suspended and dissolved substances. Such interactions give water its characteristic color and reflect information about ocean constituents, which can be captured by satellite-borne ocean color sensors in orbit hundreds of kilometers above the Earth's surface. Here, based on the spectral shape of remote sensing reflectance that is now readily available from diverse satellite ocean color sensors, the present study proposes a satellite approach for mapping surface water types in Cook Inlet, Alaska. The Iterative Self-Organizing Data Analysis Technique clustering identified 15 different water types in Cook Inlet, and then machine learning models were trained to accelerate the processing of high-resolution satellite images comprising thousands to millions of pixels each. A total of 31 classification algorithms were tested, and a neural network with medium preset showed the best performance with an accuracy of 99.9%. The main advantage of optical classification is its capability to identify water types with similar temperature or salinity properties, since the optical signatures are independent of such physical characteristics. Overall, the present study highlights that a combined approach with optical classification and the conventional temperature-salinity diagram improves our capability to differentiate water types, contributing to better monitoring of ocean dynamics that advances our baseline understanding of marine environments.

Keywords

Ocean Color Remote Sensing, Optical Classification, Machine Learning

1. Introduction

Cook Inlet, a subarctic estuary in south-central Alaska, offers a productive marine ecosystem supporting commercial and subsistence fishing. As a coastal embayment, Cook Inlet is an extremely dynamic system that responds both to variations in marine source waters and to fluctuations in freshwater input from terrestrial drainage. Marine water input to the inlet originates in the northern Gulf of Alaska. Lower salinity and elevated temperatures occur during summer owing to freshwater input and insolation [1]. Thus, Cook Inlet, together with its channels, coves, flats, and marshes, encompasses waters that are a mixture of terrestrial sources arriving from river discharge and the marine water of Shelikof Strait and the Gulf of Alaska [2].

Coastal ecosystems need to be monitored to improve our understanding of their dynamics, detect changes attributed to climate change, and assess the impact of implementing environmental protection policies [3]. Hydrographic characteristics of Cook Inlet have been investigated primarily using data collected by conventional shipboard methods [4]. More recently, nearly 90 Lagrangian surface drifters were deployed in the lower Cook Inlet and Kachemak Bay during 2003-2007, 2012, and 2017 [5]. These field campaigns substantially improved our scientific knowledge of the fundamental hydrography of Cook Inlet [5]-[8]. However, these conventional approaches are often expensive and time-consuming, making it difficult to collect large-scale, continuous time-series data.

Satellite remote sensing of ocean color offers a cost-effective, near real-time data acquisition technique that can address these limitations. Satellite ocean color observation involves the detection of spectral variations in water-leaving radiance [9], which is the sunlight backscattered out of the ocean. Ocean color is determined by the complex interactions of incident light with the optically active water constituents [10]. Generally, nearer to land, water appears greener because of increased amounts of dissolved and suspended matter [11], whereas open-ocean water appears bluer because it is very clear, somewhat similar to pure seawater.

One of the most well-established ocean color products is chlorophyll-*a* (Chl*a*), which is an important proxy for phytoplankton biomass [9]. Indeed, Chl*a* concentrations have been utilized as a key variable in studying phytoplankton dynamics at seasonal and interannual scales, to provide a better understanding of the role of the phytoplankton in diverse research fields [12]. To infer the concentration of Chl*a* or any other optically active constituents from ocean color data, algorithms have been constructed that relate the characteristics of the water signal to the property of interest [9]. Essentially, field sampling for algorithm development can be very unevenly distributed in space or time. Therefore, satellite capability may be questioned to some extent because of the uneven distribution of the field data that either forms the basis of empirical algorithms or is used for developing ocean color algorithms, raising questions regarding the applicability of the algorithms as well as the validation statistics.

To circumvent this issue, existing studies have proposed optical classification, which uses the spectrum of remote sensing reflectance ($R_{\mathrm{rs}}(\lambda)$) to quantify similarity/dissimilarity in the optical characteristics of water bodies [13]-[17]. For example, iterative refinement optical classification techniques based on $R_{rs}(\lambda)$ spectra can be powerful approaches for optical and truly homogeneous uses of ocean color remote sensing [15]. This fact suggests that monitoring ocean dynamics based on satellite-retrieved optical signatures can address the multiple shortcomings of conventional methods, including spatial and temporal resolutions, and therefore provide broader coverage in both space and time.

Here, based on optical characteristics that depend on the composition and concentration of optically active water constituents, this study presents a satellite-based mapping of optical water types in Cook Inlet. To this end, this study uses the Iterative Self-Organizing Data Analysis Technique (ISODATA) clustering technique [18] to determine optical water types from satellite-retrieved optical signatures. The ISODATA clustering technique distinguishes major optical water types by maximizing the differences between water types for a given dataset. This means it identifies optical water types differently depending on the dataset. To address this issue, we train machine-learning models using a training dataset that covers a wide range of temporal and spatial variations. The machine-learning models "know" the optical variability of the optical water types present in Cook Inlet, based on the training datasets, allowing them to retrieve optical water types accurately and consistently. Finally, the resulting optical water types are compared with those obtained from a TS diagram generated using *in situ* data to demonstrate the utility of the approach in discriminating water masses.

2. Materials and Methods

2.1. Satellite Data

Level 1A images of the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor onboard the Aqua satellite were downloaded from NASA's Ocean Color website. The MODIS/Aqua data covering over a 20-year period are the longest data record for a satellite-based global ocean color sensor. Level 3 1-km pixel resolution single-path images, containing $R_{rs}(\lambda)$ at 10 optical channels (*i.e.*, wavelengths of 412, 443, 469, 488, 531, 547, 555, 645, 667, and 678 nm) and daytime sea surface temperature (SST), were created from individual Level 1A scenes using the SeaWiFS Data Analysis System (SeaDAS) software version 9.2.0. Level 2 ocean color flags (*i.e.*, LAND, HLT, STRAYLIGHT, CLDICE) removed low-quality data during the process.

2.2. *In-Situ* Data

Sea temperature and salinity at the surface, measured with conductivity, temperature, and depth (CTD) profilers (SeaBird Electronics 19plus), were used to characterize water properties based on a TS diagram. Shipboard oceanographic surveys were made along repeated transects at generally 0.2 - 4 km spacing (**Figure 1**). These data are accessible online [19] through the DataOne repository.

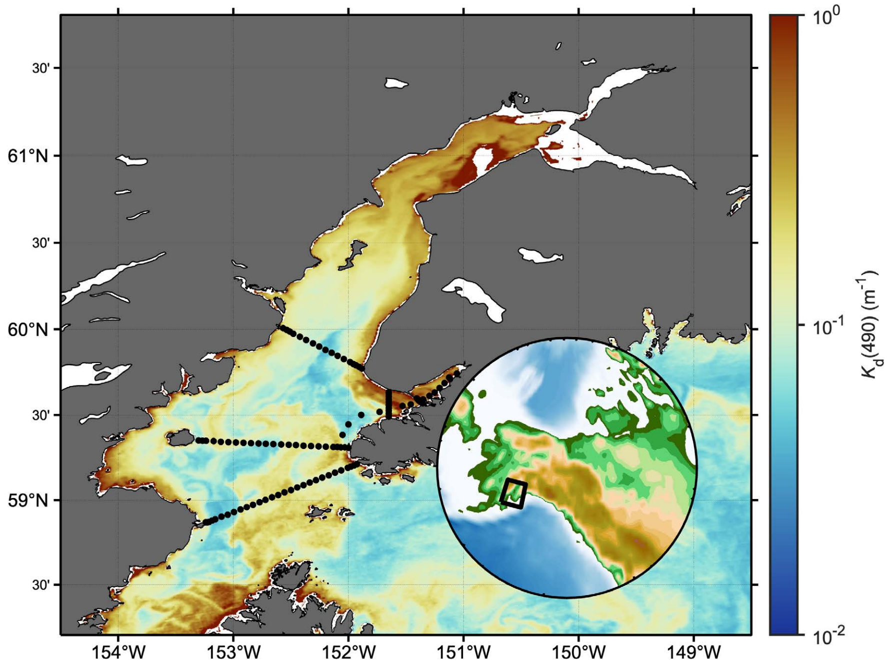


Figure 1. Map of the study area. The color represents the diffuse attenuation coefficient at 490 nm ($K_d(490)$), and the dots mark the sampling locations of the *in situ* data.

2.3. Preparing the Training Dataset

The dataset used for training machine learning models must cover the optical variability of the water masses present in Cook Inlet. To this end, this study developed a subset of the overall MODIS/Aqua data archive for 2002-2022, following the method described by [15]. Note that single-path MODIS/Aqua images were used here without compositing or averaging the images over specific time frames. To maximize the geographic and seasonal sampling of the dataset, a list of cloud-free $R_{rs}(\lambda)$ spectra was compiled for each pixel. From this list, 10 cloud-free $R_{rs}(\lambda)$ spectra per pixel (1 × 1 km) were randomly assigned to the training dataset using the MATLAB randsample function.

2.4. Identifying Optical Water Types

The classification procedures were also conducted following the method described by [15]. To reduce the first-order variability of reflectance and to focus on spectral shape, each $R_{rs}(\lambda)$ spectrum was normalized using its integrated value as follows:

$$nR_{rs}(\lambda) = R_{rs}(\lambda) \Big/ \int_{\lambda_1}^{\lambda_2} R_{rs}(\lambda)\mathrm{d}\lambda\,, \tag{1}$$

where $nR_{rs}(\lambda)$ indicates the normalized spectrum obtained by trapezoidal integration between λ_1 (412 nm) and λ_2 (678 nm).

The $nR_{rs}(\lambda)$ comprising the training dataset were then analyzed using ISODATA clustering, which is an extension of the *k*-means clustering procedure. ISODATA allows the number of clusters to be adjusted automatically during the iterative process, while *k*-means clustering assumes a fixed number of clusters.

2.5. Exploring Characteristics of Optical Water Types

This study computed major satellite ocean color products, in addition to daytime SST, to investigate the oceanographic characteristics of each water mass. Ocean color products include Chl*a*, the diffuse attenuation coefficient at 490 nm ($K_d(490)$), total suspended solids (TSS), the light absorption coefficient of colored dissolved organic matter (CDOM) at 443 nm ($a_{CDOM}(443)$), and particulate inorganic carbon (PIC). Satellite ocean color algorithms used for estimating each product are as follows: Chl*a* [20]; $K_d(490)$ [21]; TSS [22]; $a_{CDOM}(443)$ [23]; and PIC [24].

2.6. Training Machine Learning Models

The ISODATA clustering technique distinguishes optical water types by maximizing the differences between water types for a given dataset. This means that the clustering technique cannot identify unrepresented optical water types for a given dataset, even though the optical water types might be dominant in other datasets. To address this issue, the present study trained machine learning models using a training dataset that covers the optical variability of the water masses present in Cook Inlet. Additionally, a machine learning technique can accelerate processing, which represents a strong advantage when analyzing high-resolution satellite images comprising thousands to millions of pixels [25]-[27].

Machine learning is a branch of artificial intelligence that provides the ability to automatically learn and improve in inherently interdisciplinary fields based on experience without being programmed to perform specific tasks. We leveraged MATLAB's Classification Learner App in the Statistics and Machine Learning Toolbox. The app contains a wide set of 31 classification algorithms grouped into the following seven classes: tree-based, discriminant, Naive Bayes, support vector machine (SVM), K-nearest neighbors (KNN), ensemble, neural network, and kernel. Here, the training dataset consists of the optical water types (output) and the corresponding $nR_{rs}(\lambda)$ spectra (input). During training, machine learning models learn relationships between the input values and the corresponding output values without predefined or explicated equations [28]. It is noteworthy that selection of the most appropriate classifier technique is a somewhat uncertain process because the performance of the trained models relies on the dataset. Therefore, we trained multiple machine learning models using all of these classifiers, and then determined the best performing model through subsequent performance evaluation.

The training dataset comprised $nR_{rs}(\lambda)$ spectra as input features and optically derived water types as output classes. To avoid the possibility of missing certain representative samples and/or overfitting the models, a five-fold cross-validation was carried out by randomly dividing 70% of the training dataset (*i.e.*, the development subset). Evaluation of the developed models was performed five times, each time excluding one fold from the training set for testing. Each observation in the development subset was assigned to an individual group and remained in that group for the duration of the procedure so that each observation was used once for testing and four times for training the model. The final model evaluation was conducted using the remaining 30% of the training dataset (*i.e.*, the validation subset), which was completely independent of the training procedure. Finally, the best-performing model among the 31 trained models was identified based on its confusion matrix, which quantitatively summarizes the classification performance by comparing predicted and actual class labels.

2.7. Comparing Optical and TS-Based Approaches

To demonstrate the strength of our approach, this study conducted a match-up comparison [29] between optical water types determined by the optical classification and those by the conventional TS-based approach. In brief, a match-up was considered valid when *in situ* sampling of temperature and salinity was carried out within 3 hr of the satellite observation within 5 × 5 pixel boxes (5 × 5 km) centered on the position of an *in situ* observation station. The averaged $R_{rs}(\lambda)$ spectra were then used to determine the optical water types using the best-performing machine learning model. Finally, the resulting optical water type was overlaid onto the TS diagram generated using the matched *in situ* temperature and salinity at the sea surface.

3. Results and Discussion

3.1. Overview of the Training Dataset

Figure 2 shows an overview of the training dataset generated from all MODIS/Aqua images during the period 2002-2022 over Cook Inlet. The resulting dataset consists of 50,737 spectra, covering a wide range of temporal and spatial dynamics. Note that there were temporal and spatial tendencies in the number of data points assigned to the training dataset. The former is mainly due to seasonal variations in cloud coverage, whereas the latter is related to the available ocean pixels across the space.

3.2. Determination of Optical Water Types

$R_{rs}(\lambda)$ holds valuable information on the compositions and concentrations of optically active water constituents [30] and is now readily available from diverse satellite ocean color sensors. **Figure 3** shows the average $nR_{rs}(\lambda)$ spectra, together with the original $R_{rs}(\lambda)$, for each of the 15 optical water types identified in the entire training dataset. $R_{rs}(\lambda)$ shows strong variations in association with the amplitude of $R_{rs}(\lambda)$, whereas those of $nR_{rs}(\lambda)$ exhibit differences in shape, such as the locations and widths of peaks. Obviously, the original $R_{rs}(\lambda)$ showed diverse amplitudes and spectral shapes that made interpreting the data difficult. In contrast, $nR_{rs}(\lambda)$ indicated clear variations in spectral shape, primarily reflecting differences in the compositions and concentrations of optically active water constituents. In general, backscattering components determine the amplitude of $R_{rs}(\lambda)$ spectra [14], whereas that of the spectral shape is more strongly impacted by absorption components [31].

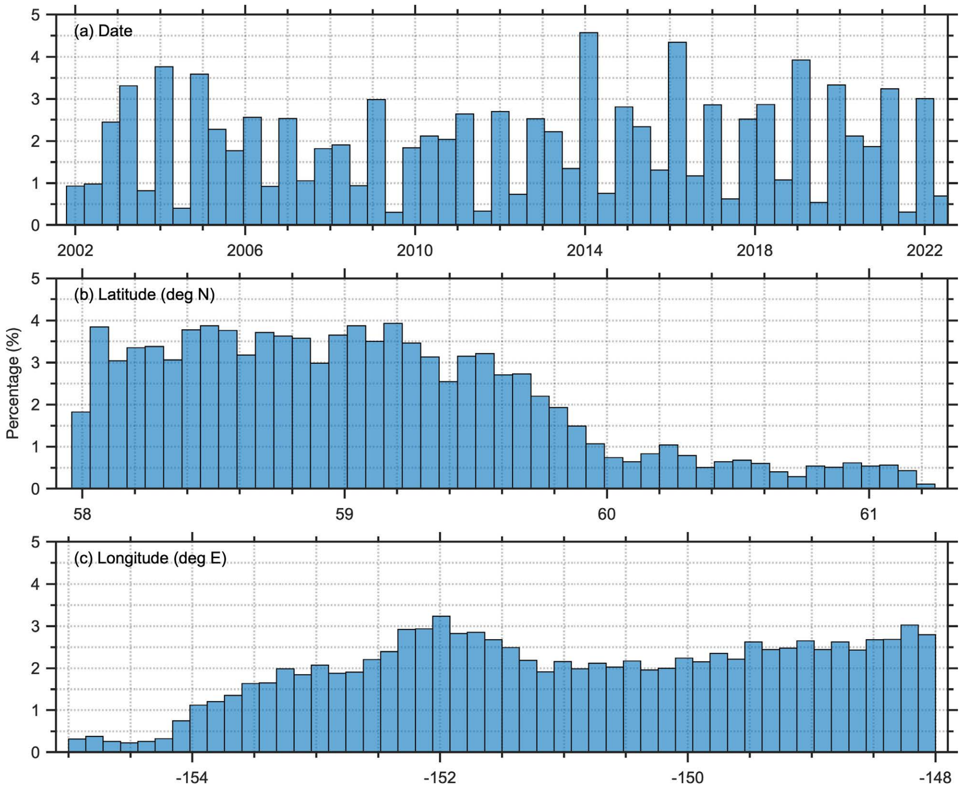


Figure 2. Overview of the training dataset. Histograms show the percentage of (a) date, (b) latitude, and (c) longitude of remote sensing reflectance ($R_{rs}(\lambda)$) spectra that were assigned to the training dataset.

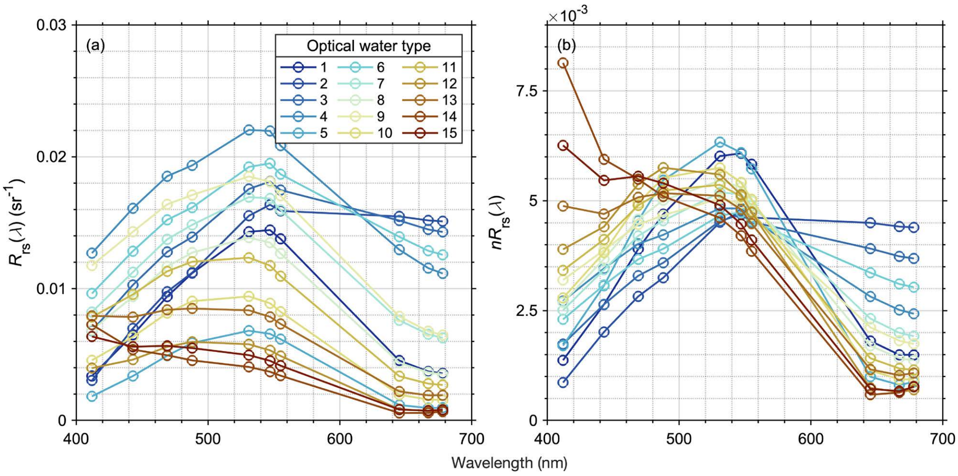


Figure 3. Optical characteristics of optical water types. Average spectra of the Moderate Resolution Imaging Spectroradiometer (MODIS) remote sensing reflectance (Rrs(λ)) at 10 wavelengths for each optical water type, determined by the Iterative Self-Organizing Data Analysis Technique (ISODATA) clustering, (a) without and (b) with normalization (nRrs(λ)).

Within the training dataset, optical water types 14 and 15, representing clear water typically in the open ocean, were not frequently observed because our target is estuarine waters around Cook Inlet (**Figure 4**). This fact indicates that optical water types distributed in limited time and space are still present in the dataset.

Optical water types 1, 5, 12, and 13 showed a similar frequency of observations in the training dataset. Other optical water types have a higher frequency, with more than 2000 observations, with the most frequent observations in optical water type 4, followed by optical water types 6, 3, and 8. Overall, our training dataset covers the diverse optical water types present in Cook Inlet, demonstrating success in maximizing seasonal and geographic coverage that benefits training machine learning models.

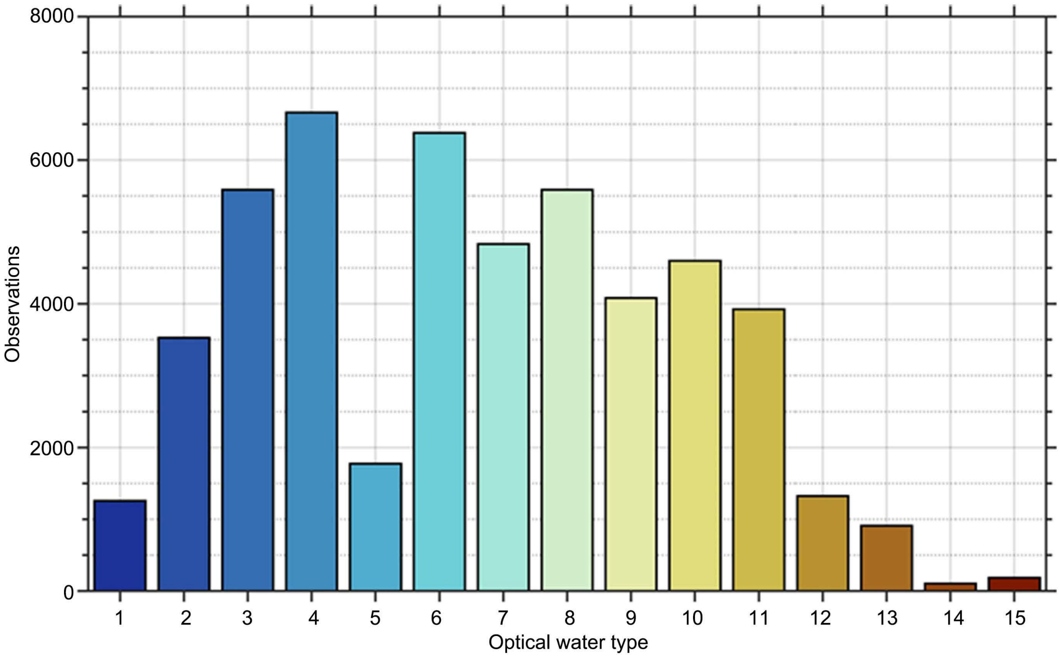


Figure 4. Frequency distribution of optical water types within the training dataset. Optical water types distributed in limited time and space are still present in the dataset, ensuring our training dataset captures diverse water types in Cook Inlet.

3.3. Oceanographic Features of Identified Optical Water Types

Major ocean color products and SST for each optical water type are shown in **Figure 5**. In general, Chl*a*, $K_d(490)$, TSS, $a_{CDOM}(443)$, and PIC varied strongly among the water types, while SST showed consistent values across them. This fact suggests that our optical classification successfully identifies different optical water types within a similar SST range, partly demonstrating the advantage of this approach in the identification of optical water types. It is noteworthy that the order of optical water types does not necessarily reflect the oceanographic features of each optical water type. Rather, the ISODATA clustering determined the order based on the similarity/dissimilarity of $nR_{rs}(\lambda)$ spectra, resulting in a random order of oceanographic features relative to the reference number of optical water types. Still, typical spectral shapes for turbid waters (*i.e.*, optical water types 1 - 6) tended to have higher $K_d(490)$ and $a_{CDOM}(443)$. Conversely, those for TSS and PIC showed random patterns, making it difficult to interpret the general oceanographic features of each optical water type.

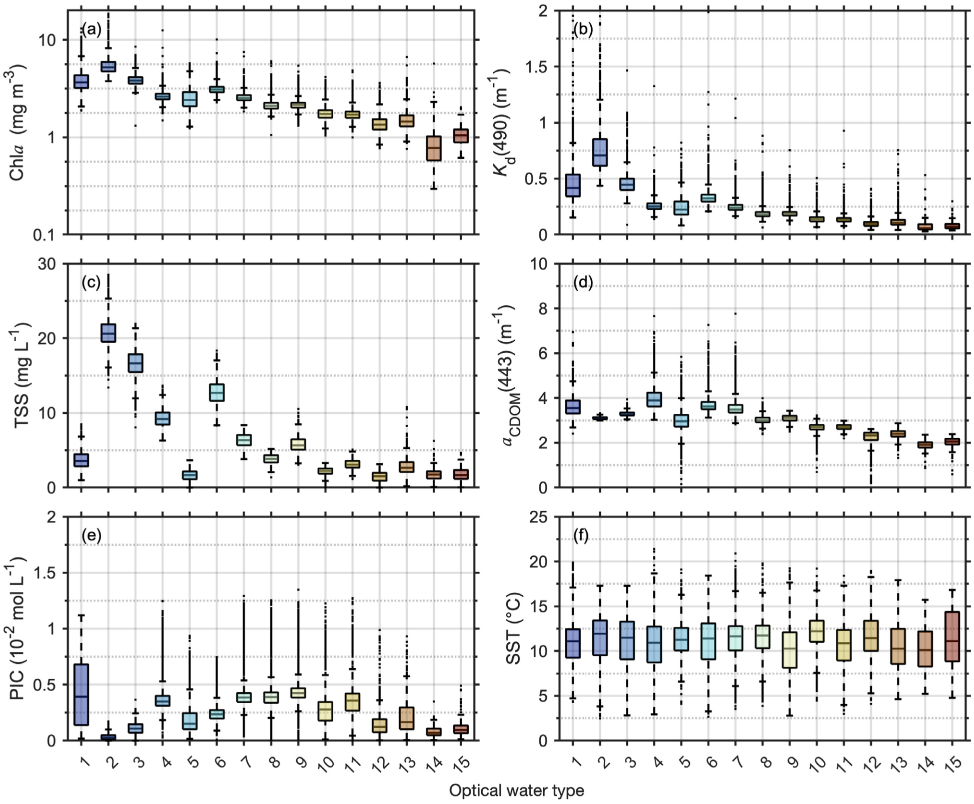


Figure 5. Oceanographic features of each optical water type: (a) Chlorophyll-*a* (Chl*a*), (b) diffuse attenuation coefficient at 490 nm ($K_d(490)$), (c) total suspended solids (TSS), (d) light absorption coefficient of colored dissolved organic matter (CDOM) at 443 nm ($a_{\text{CDOM}}(443)$), (e) particulate inorganic carbon (PIC), and (f) sea surface temperature (SST).

It is noteworthy that this study utilized existing satellite ocean color algorithms developed for the global ocean or other regions. Since the optical properties of seawater can vary among regions due to different compositions of optically active components, the performance of these ocean color algorithms may not be accurate in Cook Inlet. In particular, estuarine waters in Cook Inlet often contain extremely high volumes of suspended sediments originating from glacier meltwater entering Cook Inlet. Indeed, the concentration of suspended sediments reaches 33,750 - 67,600 mg^{-1} in Cook Inlet [32]. In comparison, estuarine waters in other regions, such as the Pearl River Delta (53 - 278 mg·L^{-1} [33]), Corpus Christi Bay (4 - 178 mg·L^{-1} [34]), and the Ganges–Brahmaputra Delta (10 - 260 mg·L^{-1} [35]), are several orders of magnitude lower than those in Cook Inlet. Since estuarine waters in Cook Inlet are typical Case-2 waters associated with the complex interplay of not only sediments but also other optically active constituents [10], careful validation and/or optimization processes need to be carried out to address the uncertainties of the existing satellite ocean color algorithms that were developed in other regions. In this context, our approach utilizing $R_{rs}(\lambda)$ instead of ocean color products has an advantage in robustness, since $R_{rs}(\lambda)$ is a radiative product that is stable regardless of the optical complexity [36]. In other words, our optical classification approach performs uniformly across a variety of water types, ranging from clear to turbid, which is a strong advantage for monitoring ocean dynamics in Cook Inlet.

3.4. Performance of Trained Machine Learning Models

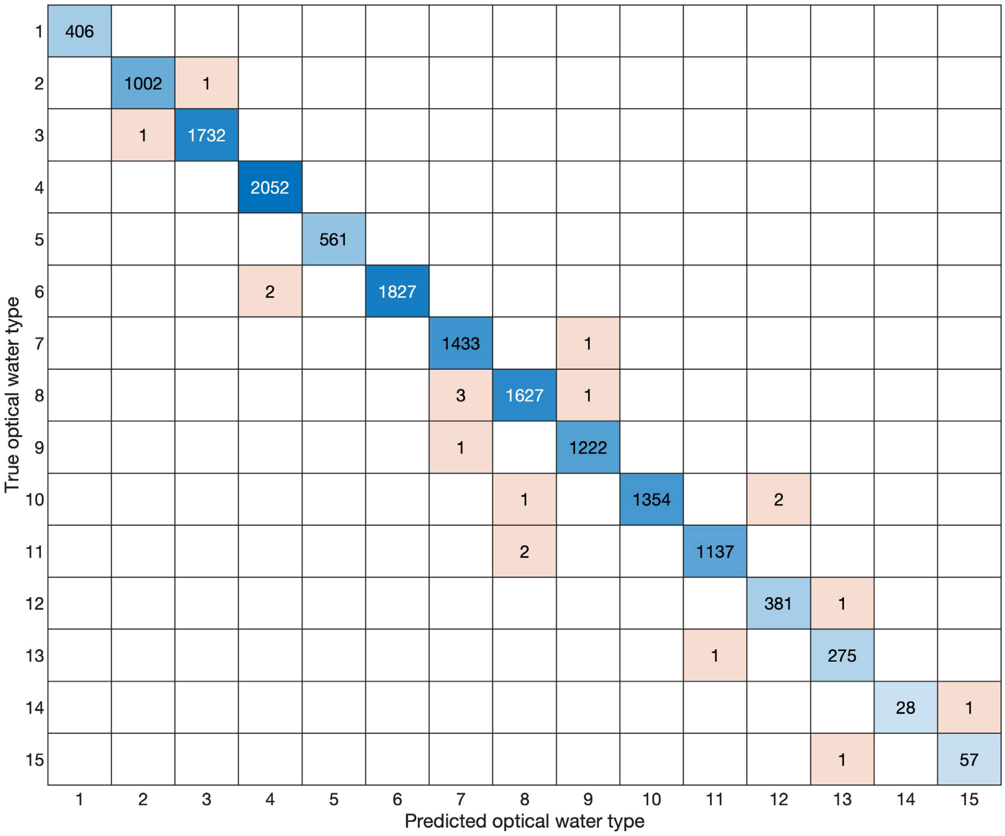
A total of 31 machine learning models were trained using the development subset ($N = 35{,}516$). The performance of each model was subsequently evaluated using the validation subset ($N = 15{,}516$) and summarized in **Table 1**. According to the statistical evaluation, there was little consistency in model performance within each classification method. For example, four out of five models trained with a neural network ranked in the top five, with the neural network with a medium preset (hereafter, MNN model) showing the best performance, with an accuracy of 99.9%, whereas all three models trained with a decision tree ranked in the bottom five, with the decision tree having a coarse preset exhibiting the worst performance, with an accuracy of 50.7%. Given that the performance of the trained models varies largely among classification methods used in the training process, it indicates that a classification method used for model training needs to be chosen carefully. The resulting confusion matrix demonstrates that the MNN model classified optical water types accurately without any bias related to the optical water types (**Figure 6**).

Table 1. Summary of the trained model's performance. The accuracy was computed as the overall percentage of the $nR_{rs}(\lambda)$ spectra of each optical water type that are correctly classified into the true class by the model in a confusion matrix based on the validation subset ($N = 35{,}516$). Abbreviations: SVM, support vector machine; KNN, K-nearest neighbors.

Model No.	Algorithm	Preset	Accuracy %
1	Tree	Fine Tree	89.9
2	Tree	Medium Tree	79.1
3	Tree	Coarse Tree	50.7
4	Discriminant	Linear Discriminant	91.1
5	Discriminant	Quadratic Discriminant	91.3
6	Naive Bayes	Gaussian Naive Bayes	92.2
7	Naive Bayes	Kernel Naive Bayes	92.9
8	SVM	Linear SVM	99.2
9	SVM	Quadratic SVM	99.1
10	SVM	Cubic SVM	98.9
11	SVM	Fine Gaussian SVM	96.2
12	SVM	Medium Gaussian SVM	98.6
13	SVM	Coarse Gaussian SVM	97.9
14	KNN	Fine KNN	95.6
15	KNN	Medium KNN	96.7
16	KNN	Coarse KNN	96.3
17	KNN	Cosine KNN	79.3

Continued

18	KNN	Cubic KNN	96.6
19	KNN	Weighted KNN	97.0
20	Ensemble	Boosted Trees	90.4
21	Ensemble	Bagged Trees	96.7
22	Ensemble	Subspace Discriminant	92.2
23	Ensemble	Subspace KNN	97.2
24	Ensemble	RUSBoosted Trees	87.2
25	Neural Network	Narrow Neural Network	99.5
26	Neural Network	Medium Neural Network	99.9
27	Neural Network	Wide Neural Network	99.6
28	Neural Network	Bilayered Neural Network	99.3
29	Neural Network	Trilayered Neural Network	98.8
30	Kernel	SVM Kernel	98.7
31	Kernel	Logistic Regression Kernel	98.5



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	406														
2		1002	1												
3		1	1732												
4				2052											
5					561										
6				2		1827									
7							1433		1						
8							3	1627	1						
9							1		1222						
10								1		1354		2			
11								2			1137				
12												381	1		
13											1		275		
14														28	1
15													1		57

Figure 6. Confusion matrix for the neural network with the medium-preset machine-learning model (MNN model). The generic position (*i*, *j*) contains a value that represents the total number of items assigned by the model to the predicted class *j* and actually belonging to the true class *i*. When *i* is equal to *j*, this value represents the percentage of examples of class *i* that are correctly classified by the algorithm. The values in the top-right and bottom-left positions of the table are very often small or zero, meaning that few $nR_{rs}(\lambda)$ spectra are classified into the wrong classes.

A major challenge of machine learning approaches is that it is difficult or impossible to derive a mechanistic understanding of the model-predicted relationship between input and output values [37]. For this reason, machine learning approaches are sometimes called "black boxes." This lack of transparency can be problematic in interpreting the results generated by the model [38] [39]. While machine learning approaches have been employed in numerous fields besides satellite remote sensing, they have not adequately addressed the issue of causality, which is essential to support wider dissemination and acceptance of the proposed models [40]. What can be said at this point is that the selection of a machine learning approach carries trade-offs between accuracy and interpretability.

As a demonstration, the MNN model was applied to a MODIS/Aqua image on September 12, 2020 (**Figure 7**). Silty waters observed in the true-color image are classified as either optical water types 5 or 12, with the relatively minor presence of some other classes within the inlet. Based on **Figure 5**, $K_d(490)$ for optical water types 5 and 12 are not so large among all 15 optical water types, suggesting that these optical water types are not turbid waters compared to others. For example, optical water types 3 and 6 are widely distributed around the mouth of Cook Inlet and the open-water areas in the Gulf of Alaska, and are characterized as turbid waters compared to optical water types 5 and 12. These results indicate an apparent mismatch between the spatial distributions of silty waters and the oceanographic characteristics of the optical water types assigned over such regions. Since silty waters should have large light attenuation associated with light absorption and scattering, optical water types 5 and 12, which correspond to silty water regions in Cook Inlet, were supposed to have larger $K_d(490)$ than those of optical water types present in the Gulf of Alaska. Overall, such mismatches between the spatial distributions of silty waters and the oceanographic characteristics of the optical water types assigned over such regions highlight uncertainties about the performance of ocean color algorithms in Cook Inlet, which is characterized as typical optically complex waters due to the prevalence of diverse optically active components.

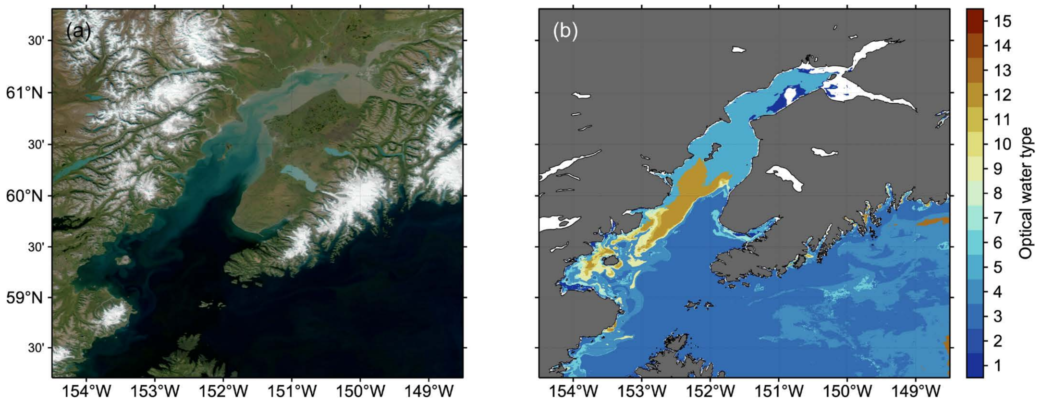


Figure 7. MODIS/Aqua images on September 12, 2020. (a) True-color image and (b) Optical water types determined by the MNN model. White pixels in panel (b) represent areas with level 2 ocean color flags.

3.5. Advantage of Optical Classification

Figure 8 shows a TS diagram for surface waters in Cook Inlet colored with optical water types determined by our optically based identification approach. Our optical classification successfully identified water types with different optical signatures even within similar temperature and salinity ranges, demonstrating superior capability for fine-scale monitoring of water mass distributions. Although this study did not explore more details of water types identified by optical signatures, different concentrations and compositions of water constituents shape distinct optical characteristics [10] and, in turn, biogeochemical features of each optical water type.

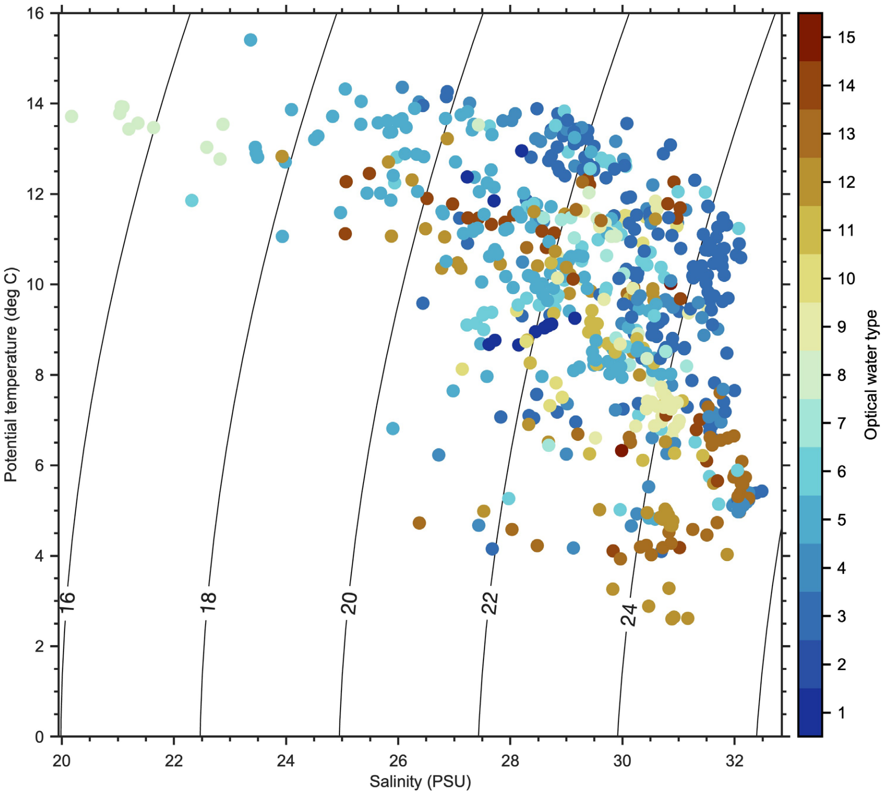


Figure 8. A temperature-salinity (TS) diagram colored by optical water types. Surface waters in Cook Inlet with valid match-ups to satellite observations are considered here.

4. Conclusion

This study proposed a satellite-based optical classification approach that can fundamentally improve our understanding of water mass distribution in Cook Inlet, thereby improving our baseline understanding of the water currents and circulation patterns in this region. The main advantage of this technique is its capability to identify water masses having similar temperature or salinity properties, since the optical signatures are independent of such physical characteristics and reflect concentrations and components of optically active constituents. Although previous studies proposed CDOM-based methods, the current work introduced an alternative optical classification method utilizing $R_{rs}(\lambda)$ spectra, circumventing considerable uncertainties in satellite ocean color products in optically complex Case-2 waters. Overall, the present study highlighted that a combined approach with optical classification and the conventional TS diagram improves our capability to identify water masses, contributing to better monitoring of ocean dynamics in Cook Inlet, a subarctic estuary with a productive marine ecosystem supporting both commercial and sport fishing activities that coexist with local wildlife.

Acknowledgements

We appreciate the Distributed Active Archive Centers at NASA's Goddard Space Flight Center (DAACs/GSFC) for the production and distribution of the ocean color data, and the lower Cook Inlet/Kachemak Bay oceanographic monitoring project for providing CTD data through the DataOne repository. Study funding was provided by the U.S. Department of the Interior, Bureau of Ocean Energy Management under Agreement Number M22AC00011. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the opinions or policies of the U.S. Government. Mention of trade names or commercial products does not constitute their endorsement by the U.S. Government.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Muench, R.D., Mofjeld, H.O. and Charnell, R.L. (1978) Oceanographic Conditions in Lower Cook Inlet: Spring and Summer 1973. *Journal of Geophysical Research*: *Oceans*, **83**, 5090-5098. https://doi.org/10.1029/jc083ic10p05090

[2] MMS (2003) Cook Inlet Planning Area, Oil and Gas Lease Sales 191 and 199: Final Environmental Impact Statement. U.S. Department of the Interior, Minerals Management Service (MMS), Alaska OCS Region, Anchorage, Alaska. https://www.boem.gov/sites/default/files/about-boem/BOEM-Regions/Alaska-Region/Environment/Environmental-Analysis/CIsV1.pdf

[3] Estes, M., Anderson, C., Appeltans, W., Bax, N., Bednaršek, N., Canonico, G., *et al.* (2021) Enhanced Monitoring of Life in the Sea Is a Critical Component of Conservation Management and Sustainable Economic Growth. *Marine Policy*, **132**, Article ID: 104699. https://doi.org/10.1016/j.marpol.2021.104699

[4] Okkonen, S.R. and Howell, S.S. (2003) Measurements of Temperature, Salinity and Circulation in Cook Inlet, Alaska. Coastal Marine Institute, University of Alaska Fairbanks. https://www.circac.org/wp-content/uploads/Okkonen-Howell-03-036.pdf

[5] Johnson, M.A. (2021) Subtidal Surface Circulation in Lower Cook Inlet and Kachemak Bay, Alaska. *Regional Studies in Marine Science*, **41**, Article ID: 101609. https://doi.org/10.1016/j.rsma.2021.101609

[6] Muench, R.D., Schumacher, J.D. and Salo, S.A. (1988) Winter Currents and Hydrographic Conditions on the Northern Central Bering Sea Shelf. *Journal of Geophysical Research*: *Oceans*, **93**, 516-526. https://doi.org/10.1029/jc093ic01p00516

[7] Oey, L., Ezer, T., Hu, C. and Muller-Karger, F.E. (2007) Baroclinic Tidal Flows and Inundation Processes in Cook Inlet, Alaska: Numerical Modeling and Satellite Observations. *Ocean Dynamics*, **57**, 205-221. https://doi.org/10.1007/s10236-007-0103-8

[8] Singhal, G., Panchang, V.G. and Nelson, J.A. (2013) Sensitivity Assessment of Wave Heights to Surface Forcing in Cook Inlet, Alaska. *Continental Shelf Research*, **63**, S50-S62. https://doi.org/10.1016/j.csr.2012.02.007

[9] Groom, S., Sathyendranath, S., Ban, Y., Bernard, S., Brewin, R., Brotas, V., *et al.* (2019) Satellite Ocean Colour: Current Status and Future Perspective. *Frontiers in Marine Science*, **6**, Article 485. https://doi.org/10.3389/fmars.2019.00485

[10] Morel, A. and Prieur, L. (1977) Analysis of Variations in Ocean Color. *Limnology and Oceanography*, **22**, 709-722. https://doi.org/10.4319/lo.1977.22.4.0709

[11] Mascarenhas, V. and Keck, T. (2018) Marine Optics and Ocean Color Remote Sensing. In: Jungblut, S., Liebich, V. and Bode, M., Eds., *YOUMARES 8—Oceans Across Boundaries*: *Learning from Each Other*, Springer International Publishing, 41-54. https://doi.org/10.1007/978-3-319-93284-2_4

[12] McClain, C.R. (2009) A Decade of Satellite Ocean Color Observations. *Annual Review of Marine Science*, **1**, 19-42. https://doi.org/10.1146/annurev.marine.010908.163650

[13] D'Alimonte, D., Melin, F., Zibordi, G. and Berthon, J. (2003) Use of the Novelty Detection Technique to Identify the Range of Applicability of Empirical Ocean Color Algorithms. *IEEE Transactions on Geoscience and Remote Sensing*, **41**, 2833-2843. https://doi.org/10.1109/tgrs.2003.818020

[14] Lubac, B. and Loisel, H. (2007) Variability and Classification of Remote Sensing Reflectance Spectra in the Eastern English Channel and Southern North Sea. *Remote Sensing of Environment*, **110**, 45-58. https://doi.org/10.1016/j.rse.2007.02.012

[15] Mélin, F. and Vantrepotte, V. (2015) How Optically Diverse Is the Coastal Ocean? *Remote Sensing of Environment*, **160**, 235-251. https://doi.org/10.1016/j.rse.2015.01.023

[16] Moore, T.S., Dowell, M.D., Bradt, S. and Ruiz Verdu, A. (2014) An Optical Water Type Framework for Selecting and Blending Retrievals from Bio-Optical Algorithms in Lakes and Coastal Waters. *Remote Sensing of Environment*, **143**, 97-111. https://doi.org/10.1016/j.rse.2013.11.021

[17] Vantrepotte, V., Loisel, H., Dessailly, D. and Mériaux, X. (2012) Optical Classification of Contrasted Coastal Waters. *Remote Sensing of Environment*, **123**, 306-323. https://doi.org/10.1016/j.rse.2012.03.004

[18] Tou, J.T. and Gonzalez, R.C. (1974) Pattern Recognition Principles. NASA.

[19] Holderied, K. and Renner, M. (2023) Oceanographic Monitoring in Cook Inlet and Kachemak Bay, CTD Data, 2012-2022, Gulf Watch Alaska Environmental Drivers Component. https://doi.org/10.24431/rw1k32v

[20] O'Reilly, J.E. and Werdell, P.J. (2019) Chlorophyll Algorithms for Ocean Color Sensors—OC4, OC5 & Oc6. *Remote Sensing of Environment*, **229**, 32-47. https://doi.org/10.1016/j.rse.2019.04.021

[21] Werdell, P.J. and Bailey, S.W. (2005) An Improved *In-Situ* Bio-Optical Data Set for Ocean Color Algorithm Development and Satellite Data Product Validation. *Remote Sensing of Environment*, **98**, 122-140. https://doi.org/10.1016/j.rse.2005.07.001

[22] Miller, R.L. and McKee, B.A. (2004) Using MODIS Terra 250 M Imagery to Map Concentrations of Total Suspended Matter in Coastal Waters. *Remote Sensing of En-*

vironment, **93**, 259-266. https://doi.org/10.1016/j.rse.2004.07.012

[23] Lee, Z., Carder, K.L. and Arnone, R.A. (2002) Deriving Inherent Optical Properties from Water Color: A Multiband Quasi-Analytical Algorithm for Optically Deep Waters. *Applied Optics*, **41**, 5755-5772. https://doi.org/10.1364/ao.41.005755

[24] Balch, W.M., Gordon, H.R., Bowler, B.C., Drapeau, D.T. and Booth, E.S. (2005) Calcium Carbonate Measurements in the Surface Global Ocean Based on Moderate-Resolution Imaging Spectroradiometer Data. *Journal of Geophysical Research*: *Oceans*, **110**, C07001. https://doi.org/10.1029/2004jc002560

[25] Paul, S. and Huntemann, M. (2021) Improved Machine-Learning-Based Open-Water—Sea-Ice—Cloud Discrimination over Wintertime Antarctic Sea Ice Using MODIS Thermal-Infrared Imagery. *The Cryosphere*, **15**, 1551-1565. https://doi.org/10.5194/tc-15-1551-2021

[26] Rösel, A., Kaleschke, L. and Birnbaum, G. (2012) Melt Ponds on Arctic Sea Ice Determined from MODIS Satellite Data Using an Artificial Neural Network. *The Cryosphere*, **6**, 431-446. https://doi.org/10.5194/tc-6-431-2012

[27] Waga, H., Eicken, H., Light, B. and Fukamachi, Y. (2022) A Neural Network-Based Method for Satellite-Based Mapping of Sediment-Laden Sea Ice in the Arctic. *Remote Sensing of Environment*, **270**, Article ID: 112861. https://doi.org/10.1016/j.rse.2021.112861

[28] Marzban, C. (2009) Basic Statistics and Basic AI: Neural Networks. In: Haupt, S.E., Pasini, A. and Marzban, C., Eds., *Artificial Intelligence Methods in the Environmental Sciences*, Springer, 15-47. https://doi.org/10.1007/978-1-4020-9119-3_2

[29] Bailey, S.W. and Werdell, P.J. (2006) A Multi-Sensor Approach for the On-Orbit Validation of Ocean Color Satellite Data Products. *Remote Sensing of Environment*, **102**, 12-23. https://doi.org/10.1016/j.rse.2006.01.015

[30] Gordon, H.R., Brown, O.B., Evans, R.H., Brown, J.W., Smith, R.C., Baker, K.S., *et al.* (1988) A Semianalytic Radiance Model of Ocean Color. *Journal of Geophysical Research*: *Atmospheres*, **93**, 10909-10924. https://doi.org/10.1029/jd093id09p10909

[31] Sathyendranath, S., Platt, T., Caverhill, C.M., Warnock, R.E. and Lewis, M.R. (1989) Remote Sensing of Oceanic Primary Production: Computations Using a Spectral Model. *Deep Sea Research Part A. Oceanographic Research Papers*, **36**, 431-453. https://doi.org/10.1016/0198-0149(89)90046-0

[32] Wagner, N. (2021) Cook Inlet Sediment Budget and Water Quality Model. University of Alaska Anchorage. https://scholarworks.alaska.edu/handle/11122/12021

[33] Cao, B., Qiu, J., Zhang, W., Xie, X., Lu, X., Yang, X., *et al.* (2022) Retrieval of Suspended Sediment Concentrations in the Pearl River Estuary Using Multi-Source Satellite Imagery. *Remote Sensing*, **14**, Article 3896. https://doi.org/10.3390/rs14163896

[34] Reisinger, A., Gibeaut, J.C. and Tissot, P.E. (2017) Estuarine Suspended Sediment Dynamics: Observations Derived from over a Decade of Satellite Data. *Frontiers in Marine Science*, **4**, Article 233. https://doi.org/10.3389/fmars.2017.00233

[35] Tilstone, G.H., Angel-Benavides, I.M., Pradhan, Y., Shutler, J.D., Groom, S. and Sathyendranath, S. (2011) An Assessment of Chlorophyll-A Algorithms Available for Seawifs in Coastal and Open Areas of the Bay of Bengal and Arabian Sea. *Remote Sensing of Environment*, **115**, 2277-2291. https://doi.org/10.1016/j.rse.2011.04.028

[36] Mobley, C.D. (1994) Light and Water. Academic Press.

[37] Ray, S. (2019) A Quick Review of Machine Learning Algorithms. 2019 *International Conference on Machine Learning*, *Big Data*, *Cloud and Parallel Computing* (*COMITCon*), Faridabad, 14-16 February 2019, 35-39.

https://doi.org/10.1109/comitcon.2019.8862451

[38] Vollmer, S., Mateen, B.A., Bohner, G., Király, F.J., Ghani, R., Jonsson, P., *et al.* (2020) Machine Learning and Artificial Intelligence Research for Patient Benefit: 20 Critical Questions on Transparency, Replicability, Ethics, and Effectiveness. *BMJ*, **368**, L6927. https://doi.org/10.1136/bmj.l6927

[39] Wachter, S., Mittelstadt, B. and Russell, C. (2017) Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR. *Harvard Journal of Law & Technology*, 31. https://doi.org/10.2139/ssrn.3063289

[40] Hall, O., Ohlsson, M. and Rögnvaldsson, T. (2022) A Review of Explainable AI in the Satellite Data, Deep Machine Learning, and Human Poverty Domain. *Patterns*, **3**, Article ID: 100600. https://doi.org/10.1016/j.patter.2022.100600