ABSTRACT

Title of Dissertation:	INCORPORATING UNOCCUPIED AIRCRAFT SYSTEMS (UAS) AND EARTH OBSERVING SATELLITES TO ENHANCE ENVIRONMENTAL REMOTE SENSING OF CHESAPEAKE BAY
	Anna Elizabeth Windle, Doctor of Philosophy, 2023
Dissertation directed by:	Dr. Greg Silsbe Marine Estuarine Environmental Sciences

Environmental remote sensing is the science of monitoring physical, chemical, and biological characteristics of the Earth through space and time, and from a distance, by measuring how these environments interact with electromagnetic energy, or more simply through changes in color. This dissertation leverages *in situ*, satellite, and unoccupied aircraft system (UAS, drones) data to enhance the efficacy of environmental remote sensing in Chesapeake Bay. Satellite data consists of distinct contributions of the surface under observation and the intervening atmosphere. Atmospheric correction (AC) processors seek to isolate the surface signal, and while several variants exist, their accuracy varies widely in optically complex coastal waters. Chapter 2 is a statistical evaluation of four common AC variants applied to data collected by the most recent operational ocean color sensor, the Ocean Land Color Instrument (OLCI) onboard Copernicus Sentinel-3A and -3B satellites. Remote

sensing reflectance (R_{rs}), the product of AC processors from which a suite of water quality metrics is then derived, was obtained from each AC variant and matched in space and time with *in situ* R_{rs} data collected in the Chesapeake Bay. AC results varied widely, and the most statistically robust was a neural-net based algorithm (Case 2 Regional Coast Color, C2RCC). The resultant shape and magnitude of R_{rs} (e.g., color) is governed by the type and concentration of optically active constituents (OACs), namely phytoplankton pigments, chromophoric dissolved organic matter, inorganic sediment, and water itself. In coastal waters where OACs are dynamic and vary independently from each other, deriving accurate water quality metrics remains an open challenge. Chapter 3 applies a spectral clustering classification of OLCI R_{rs} data (2016-2022) and identifies the fifteen most dominant optical water types (OWTs) of Chesapeake Bay. OWTs were matched in space and time with Chesapeake Bay water quality monitoring data, and a statistical evaluation demonstrates how water quality data are constrained within and across OWTs. In contrast to earth-observing satellites, UAS equipped with optical sensors offer on-demand, highly resolved data. Aquatic UAS applications are in their infancy, and the critical removal of light reflected directly off the skin of water has received little attention in the literature. Chapter 4 proposes four different approaches to remove direct surface reflectance from UAS imagery and evaluates each against *in situ* R_{rs} data. The most accurate method is a simple empirical model that exploits measurements in the infrared where water strongly absorbs light; applying this model permits high resolution water quality retrievals with only modest uncertainty. Chapter 5 uses UAS imagery to monitor a wetland restoration site in the Chesapeake Bay across seasons and years. A supervised random forest model is developed with UAS data and used to classify species-specific marsh vegetation with 97-99% accuracy. Vegetation classification maps were compared to as-built planting plans to delineate instances of significant marsh migration. Chapter 6 summarizes how the environmental remote sensing methods used in this dissertation can contribute to a better understanding of coastal research, monitoring, and management by addressing challenges, gaps, and potential solutions at various scales.

Author's Note

The chapters that compromise this dissertation were written in manuscript format for scientific publication. Thus, the formatting of each chapter follows the guidelines of the journal to which the manuscript was submitted. Citations for individual chapters are as follows:

Chapter 2: Windle, A.E., Evers-King, H., Loveday, B.R., Ondrusek, M. and Silsbe, G.M., 2022. Evaluating atmospheric correction algorithms applied to OLCI Sentinel-3 data of Chesapeake Bay Waters. *Remote Sensing*, *14*(8), p.1881. https://doi.org/10.3390/rs14081881

Chapter 4: Windle, A.E. and Silsbe, G.M., 2021. Evaluation of Unoccupied Aircraft System (UAS) remote sensing reflectance retrievals for water quality monitoring in coastal waters. *Frontiers in Environmental Science*, *9*, p.674247. https://doi.org/10.3389/fenvs.2021.674247

Chapter 5: Windle, A.E., Staver, L., Elmore, A.J., Scherer, S., Keller, S., Malmgren, B. and Silsbe, G.M., 2023. Multi-temporal high-resolution marsh vegetation mapping using Unoccupied Aircraft System remote sensing and machine learning. *Frontiers in Remote Sensing*, 4, p.17. https://doi.org/10.3389/frsen.2023.1140999

INCORPORATING UNOCCUPIED AIRCRAFT SYSTEMS (UAS) AND EARTH OBSERVING SATELLITES TO ENHANCE ENVIRONMENTAL REMOTE SENSING OF CHESAPEAKE BAY

by

Anna Elizabeth Windle

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2023

Advisory Committee:
Dr. Greg Silsbe, Chair
Dr. Andrew Elmore
Dr. Judith O'Neil
Dr. Raleigh Hood
Dr. Larry Sanford
Dr. Dongdong Wang, Dean's Representative

Acknowledgements

As I look back on the last five years, I feel gratitude, nostalgia, and joy. My time at Horn Point Lab has been tremendously influential to my career and who I am as a scientist and person. There are many people to acknowledge that have educated, supported, and encouraged me.

My advisor, Dr. Greg Silsbe for allowing me to explore my research ideas to their fullest potential, providing me the upmost support and encouragement, and always being there when I needed assistance. I will strive to be a mentor to others as you have been to me.

My committee members, who have all significantly contributed to my research and served as influential scientists in my life.

My collaborators for providing guidance, technical support, and help in the field and lab.

The Horn Point Lab community, which has been my family for the past five years. I am honored and grateful to have been supported by a Horn Point Lab assistantship to conduct my research. Every person at the lab has made an impact on my life and I will cherish the memories I have made here.

And finally, my family and friends, who have endlessly supported me. My parents have backed all my educational and career decisions and given me endless amounts of encouragement. Mom, thank you for being a prominent role model throughout my life and always giving me confidence that I can do whatever I put my mind to. Joe, thank you for everything. You are a light in my life, and I greatly appreciate you.

Table of Contents

List of Tables

Chapter 4

Chapter 5

List of Figures

Figure 2. A) MicaSense RedEdge-MX multispectral sensor and downwelling light sensor (DLS) collects total radiance (L_T) (sum of water-leaving radiance, L_w and surface reflected radiance, L_{SR}) and downwelling irradiance (E_d) measurements while in a low altitude flight. Sky radiance (L_{sky}) is collected by positioning the sensor at a 40° angle from zenith away from sun. B) MicaSense RedEdge-MX multispectral sensor and DLS collects L_T and E_d in 5 wavebands: red, green, blue, red-edge, near-infrared. C) Approximate spectral response functions of MicaSense RedEdge-MX sensor.

Figure 3. Visualization of the three-dimensional ρ look-up table derived from HydroLight simulations with varying solar zenith angles, wind speed (m/s), and cloud cover (%) corresponding to a nadir viewing angle.

Figure 4. Nonlinear relationship between a blue (475 nm) to red edge (717 nm) ratio of total UAS reflectance to *in situ* R_{rs} in the NIR band (842 nm). An exponential model was used to estimate UAS-derived R_{rs}(NIR) baseline in order to estimate L_{SR}.

Figure 5. A) *In situ* remote sensing reflectance (R_{rs}) spectra collected with hyperspectral TriOS radiometers and B) Sky radiance (L_{sky}) spectra collected with MicaSense RedEdge-MX multispectral camera at stations along the Choptank River on 9/16/20 (blue), 10/01/20 (brown), and 10/22/20 (green).

Figure 6. Example of an individual UAS image (green band) with different radiometric values: (A) R_{UAS}, (B) R_{UAS} with initial sun glint masking and (C-F) remote sensing reflectance (R_{rs}) using various methods to remove surface reflected light: (C) ρ look-up table (LUT), (D) Dark pixel assumption with NIR = 0, E) Dark pixel assumption with NIR > 0, F) Deglingting methods following Hochberg et al. (2003).

Figure 7. Total UAS reflectance (R_{UAS}) (A) and remote sensing reflectance (R_{rs}) spectra using various methods to remove surface reflected light: B) ρ look-up table from HydroLight simulations, C) Dark pixel assumption with NIR = 0, D) Dark pixel assumption with NIR > 0, E) Deglinting methods following Hochberg et al. (2003), and F) *In situ* R_{rs} spectra from TriOS sensors with MicaSense SRFs applied. Negative values are not shown in plots.

Figure 8. Comparison of UAS radiometry and *in situ* R_{rs} in all bands at each station (n=28) using different methods to remove surface reflected light after initial sun glint masking. A) Total UAS-derived reflectance (R_{UAS}), B) ρ look-up table from HydroLight simulations, C) Dark pixel assumption with NIR = 0, D) Dark pixel assumption with NIR > 0, E) Deglinting methods following Hochberg et al. (2003). Negative values are not shown in plots.

Figure 9. Comparison between *in situ* and modelled A) chlorophyll *a* concentration and B) TSS concentration from multiple linear regressions.

model without each variable, where a higher value indicates greater importance of that variable in predicting vegetation.

Figure 6. Bar plot of random forest classification accuracies (%) averaged across datasets when applied to various combinations of raster layers. OA = Overall accuracy, UA = User’s accuracy, PA = Producer’s accuracy.

Figure 7. Field photos capturing spectral and structural differences between *Spartina alterniflora* and *Spartina patens* during the (A) fall and (B) summer seasons.

Figure S1. (A) Ground control point (GCP) deployed in the marsh and (B) surveying the GCP with an RTK GPS.

Figure S2. Scatterplot of elevation measurements extracted from the UAS-SfM derived DTM (open circles) and DSM (open triangles) at the location of RTK GPS elevation measurements. Elevation measurements are colored black, blue, green, and red for the August 2020, April 2021, June 2022, and November 2022 datasets, respectively. RTK GPS checkpoints were not collected in November 2019. Black line is the 1:1 line.

Chapter 1: Introduction

Introduction to environmental remote sensing

All matter on Earth reflects, absorbs, and transmits electromagnetic energy, the degree to which varies with wavelength. Environmental remote sensing is the science of obtaining information at a distance by measuring electromagnetic energy, either from the sun (passive) or an instrument (active), reflected off matter in the Earth's biosphere. Remote platforms such as earth observing satellites, aircraft, and unoccupied aircraft systems (UAS, or drones) can carry instruments that measure the intensity of reflected energy, defined as radiance, at select wavelengths that together constitute a spectrum. This spectral data can complement *in situ* field observations, allowing for extensive, synoptic, and repeated monitoring of the atmosphere, terrestrial environments, and the surface mixing layer of aquatic environments.

Satellite remote sensing has become one of the most efficient tools to survey the Earth at local, regional, and global spatial scales, albeit with some challenges due to the intervening atmosphere. Since the radiation source (i.e., the sun) and the satellite detector are both 'remote', the radiance incident on a given sensor is affected not only by the reflection of light from the surface, but also by interactions with the atmosphere. Over dark water surfaces, it is widely recognized that 90–99% of the total signal measured by a satellite sensor is due to confounding influences from the atmosphere and varying water surface facets (IOCCG, 2010). Accurately quantifying and ultimately removing these influences, termed atmospheric correction, enables the science of satellite remote sensing. Coastal waters in close proximity to land are also susceptible to the adjacency effect where radiance from high reflective areas spill over neighboring low reflective regions, modifying their apparent brightness (Bulgarelli et al., 2014). Pixel flagging techniques can be applied during atmospheric correction processing to mask pixels contaminated by land adjacency effects (Feng & Hu, 2017).

Suborbital remote sensing capabilities such as UAS remote sensing can also extend *in situ* data collection, collecting high resolution data not amenable to space-based observations (Anderson & Gaston, 2013; Johnston, 2019). The range of UAS aircraft (i.e., fixed-wing, rotary-wing) and the selection of modular sensors (e.g., multispectral, hyperspectral, lidar, etc.) make UAS an adaptable tool to for accessing and remote sensing ecosystems with on-demand capabilities without the challenges of cloud cover, atmospheric, and land adjacency effects associated with satellite and airborne remote sensing. While technical and regulatory hurdles exist, significant progress has been made within the last decade to utilize UAS remote sensing for responsive, timely, and cost-effective monitoring of environmental phenomena at scales relevant for many ecologically relevant variables (Johnston, 2019).

Coastal environments, at the nexus of land and water, are highly dynamic spaces where physical, chemical, and biological processes take place across a range of temporal and spatial scales. Furthermore, over one-third of the total human population live within 100 km of an oceanic coast, leading to a close connection with anthropogenic stressors (IOCCG, 2008). Consequently, coastal environments are also some of the most productive, diverse, and rapidly changing ecosystems on Earth. Coastal environments have already begun to experience lasting shifts in community composition, productivity, and biodiversity from a changing climate, making it essential that services and values of these dynamic systems are safeguarded and sustained for future generations (Doney et al., 2011). As technology and knowledge continues to advance, environmental remote sensing will continue to serve as a powerful tool to collect data at previously unobtainable temporal and spatial scales (Ridge & Johnston, 2020)

This dissertation demonstrates how earth observing satellites and unoccupied aircraft systems (UAS) can be incorporated to enhance environmental remote sensing of coastal environments. The objectives of this introductory chapter are to explain the basic theory of environmental remote sensing, provide relevant background information on environmental remote sensing in the context of the Chesapeake Bay estuarine system, and summarize the specific remote sensing techniques used in this dissertation.

Theory

Remote sensors measure radiance as the radiant flux per unit of solid angle, which corresponds to the intensity of electromagnetic radiation in a given direction towards the sensor. Since radiance is dependent on the illumination of the observed area, it is normalized by hemispherical downwelling solar irradiance to calculate surface reflectance, that is the amount of radiation a surface reflects per unit of radiation incident upon it. After additional atmospheric correction processing, surface reflectance is computed and serves as the primary input for terrestrial geophysical parameters, contributing to land cover classifications, vegetation indices, albedo, and more (Vermote et al., 2008; Roy et al., 2014).

In aquatic environments, solar radiation enters the surface of water and interacts with in-water optically active constituents (OACs) such as phytoplankton pigments, colored dissolved organic matter (CDOM), and suspended inorganic particles (Kirk, 1994). Therefore, the amount of radiance that exits the water and collected by a satellite sensor is altered by the type and concentrations of OACs. This quantity, termed water leaving radiance, is atmospherically corrected and normalized by hemispherical downwelling solar irradiance to obtain remote sensing reflectance (R_{rs}). In theory, R_{rs} is to a first approximation, proportional to the ratio of backscattering to absorption properties of OACs. Accordingly, the amount of absorption and backscattering occurring in the water determines the magnitude and spectral shape of R_{rs}, which ultimately illustrates the apparent color and brightness of water bodies. R_{rs} is the fundamental quantity to derive aquatic bio-optic and geophysical parameters such as chlorophyll *a* concentrations, total suspended matter concentrations, diffuse attenuation, and more (IOCCG, 2006).

In open ocean waters, there are typically low concentrations of OACs and R_{rs} is dominated by pure water absorption; the spectrum is characterized by a peak in the blue portion of the visible (Figure 1). In most estuaries, absorption by CDOM and phytoplankton pigments combined with absorption and scattering of suspended particles shifts the peak R_{rs} into the green and red portions of the visible spectrum with secondary peaks in the red from phytoplankton pigments and fluorescence (Figure 1). In sediment-rich waters, R_{rs} tends to increase in magnitude due to the highly scattering properties of inorganic particles and peaks in the yellow to red wavelengths due to absorption of non-algal particles (e.g., detritus, inorganic minerals) (Figure 1). The Chesapeake Bay is optically complex, owing to OACs that vary independently from one another. In general, water quality in the Bay follows a latitudinal trophic gradient with the fresh Upper Bay typically being more eutrophic, containing higher concentrations of OACs and the saline Lower Bay region being more mesotrophic with lower concentrations of OACs.

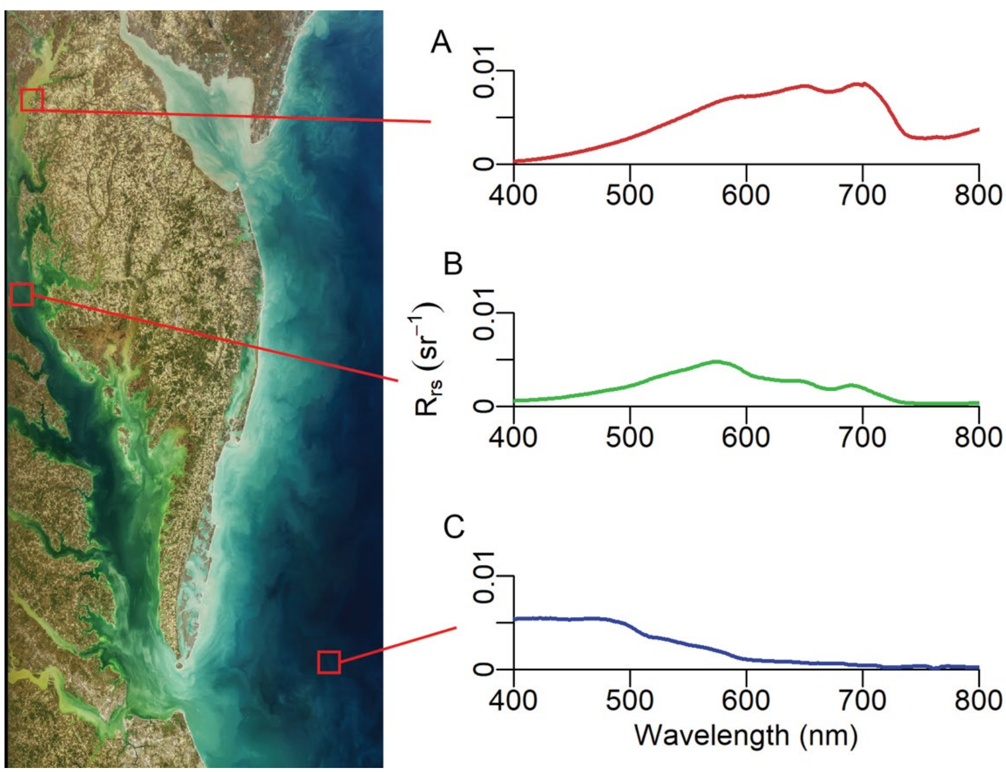


Figure 1. *In situ* hyperspectral R_{rs} measurements representative of A) estuarine waters, B) sediment-rich waters, and C) open ocean waters.

History of remote sensing in the Chesapeake Bay

The Chesapeake Bay has a long history of remote sensing research. Before earth observing satellites, imagery collected from aircraft was used to assess water quality, identify sources of water pollution, and map wetland species composition (Anderson, 1970). Since 1984, aircraft remote sensing has been used to annually monitor changes in submerged aquatic vegetation (SAV) coverage (Orth et al., 2017). Furthermore, from 1989 to 2013, the Chesapeake Bay Remote Sensing Program used aircraft remote sensing to estimate chlorophyll *a* concentrations and study interannual variability of phytoplankton distributions in the Bay (Harding et al., 1992; Miller & Harding, 2007).

After the first Landsat 1 satellite launched in 1972, satellite imagery was used to study many environmental phenomena in the Bay such as sediment plume dynamics and transport (Shelley, 1976; Munday & Fedosh, 1981; Stumpf, 1988), submerged aquatic vegetation (Ackleson & Klemas, 1987), coastal wetlands (Anderson et al., 1975), and water quality (Fleischer et al., 1976). After the launch of the first proof of concept ocean color satellite in 1978, the Bay served as a testbed for many 'ocean color' remote sensing studies (Magnuson et al., 2004; Harding et al., 2005; Acker et al., 2005). These studies attempted to quantify water quality properties by applying locally tuned empirical band ratio algorithms (Harding et al., 2005; Gitelson et al., 2007; Werdell et al., 2009). However, due to spatial variability in

OACs, empirical tuning using geographically limited *in situ* data can fail in areas with complex optical properties (Gitelson et al., 2007). Consequently, modified empirical and physics-based (analytical) algorithms have been tested on the Chesapeake Bay to improve water quality retrievals (Tzortziou et al., 2007; Son & Wang et al., 2012; Werdell et al., 2009; Ondrusek et al., 2012; Le et al., 2013; Zheng & DiGiacomo, 2017; Gilerson et al., 2021). Follow-on satellite ocean color instruments and many advanced water quality retrieval algorithms have been applied in Chesapeake Bay in an attempt to improve optically complex water quality retrievals for enhanced monitoring and management (Franz et al., 2015; Tomlinson et al., 2019; Wolny et al., 2020; Pahlevan et al., 2020).

Dissertation outline

This dissertation demonstrates how a combination of *in situ*, satellite, and UAS data can enhance environmental remote sensing in Chesapeake Bay, the largest estuary in the continental U.S. While many aquatic satellite remote sensing studies have been conducted in Chesapeake Bay, no single atmospheric correction processor has emerged as an optimal method. The first research chapter of this dissertation (Chapter 2) uses a regional *in situ* dataset from the Chesapeake Bay to evaluate the performance of four atmospheric correction algorithms applied to satellite data collected by the Ocean Land Color Instrument (OLCI) onboard the Copernicus Sentinel-3A and -3B satellites. OLCI currently contains the most spectral bands and the highest spatial resolution (300 m) of the current suite of missions designed for ocean color. The constellation of OLCI also provides nearly daily coverage of Chesapeake Bay, making it an ideal sensor to study complex, small-scale dynamic processes typical of a coastal water body. The results recommend a well performing atmospheric correction algorithm to use for Chesapeake Bay waters and provides a framework with associated uncertainties and recommendations to utilize OLCI ocean color data to monitor the water quality and biogeochemical dynamics in Chesapeake Bay.

Understanding the regional variability in estuarine dynamics has historically been confined to ship-board surveys and mooring observations, which are limited in time and space. The second research chapter (Chapter 3) synthesizes water quality dynamics in the Chesapeake Bay by applying a machine learning technique to classify remote sensing data into optical water types (Chapter 3). Accompanying *in situ* data are analyzed within the statistically defined water types to study biogeochemical properties throughout space and time. The results of this study aid in the interpretation of Bay-wide water quality trends, assist in the dynamic selection of water quality retrieval algorithms, and support identifying regions of water quality management concern.

The last two research chapters focus on UAS remote sensing that offers scale-appropriate measurements of coastal environmental phenomena. A UAS equipped with off-the-shelf optical sensors originally designed for terrestrial applications can be used to derive water quality properties in coastal waters. Since the methods to measure water reflectance from a UAS have not been widely considered, some studies have erroneously estimated water quality retrievals. The at-sensor total radiance a UAS measured constitutes the sum of water-leaving radiance and incident radiance reflected off the sea surface into the detector's field of view. Failure to accurately account for surface reflected light can significantly influence remote sensing reflectance, resulting in inaccurate water quality estimates once algorithms are applied. The third research chapter (Chapter 4) evaluates the efficacy of methods that remove surface reflected light from total UAS radiance measurements in order to derive more accurate remotely sensed retrievals of scientifically valuable in-water constituents. The results recommend a method that exploits the high absorption of water at near-infrared wavelengths to estimate and remove surface reflected light and provide further insight into using UAS remote sensing to improve high resolution water quality monitoring of coastal water bodies.

Monitoring ecosystem restoration can be labor intensive, short-term, and potentially harmful to a habitat. While satellite remote sensing is a valuable monitoring tool; the spatial and temporal resolution of imagery often places operational constraints, especially in small or spatially complex environments. The last research chapter (Chapter 5) uses UAS remote sensing to monitor a wetland restoration site in the Chesapeake Bay. UAS data was used as input into supervised random forest classification models to classify species-specific marsh vegetation. High resolution vegetation classification maps were compared to as-built planting plans, demonstrating significant changes in vegetation and potential instances of marsh migration. The results of this chapter can be used by scientists and managers to reliably monitor and assess marsh migration and restoration success in a changing climate.

The conclusions chapter (Chapter 6) summarizes the major findings of each research chapter and describes how environmental remote sensing methods used in this dissertation can contribute to a better understanding of coastal research, monitoring, and management by addressing challenges, gaps, and potential solutions at various scales. Also included is a description of the strengths and weaknesses between *in situ*, satellite, and UAS remote sensing and considerations on selecting the best method for a research question.

References

Acker, J.G., Harding, L.W., Leptoukh, G., Zhu, T. and Shen, S., (2005). Remotely-sensed chlorophyll *a* at the Chesapeake Bay mouth is correlated with annual freshwater flow to Chesapeake Bay. *Geophysical Research Letters*, *32*(5).

Ackleson, S.G. and Klemas, V., (1987). Remote sensing of submerged aquatic vegetation in lower Chesapeake Bay: A comparison of Landsat MSS to TM imagery. *Remote Sensing of Environment*, *22*(2), pp.235-248.

Anderson, R., Alsid, L. and Carter, V., (1975), June. Comparative utility of Landsat-1 and Skylab data for coastal wetland mapping and ecological studies. In *NASA. Lyndon B. Johnson Space Center NASA Earth Resources Survey Symp. Vol. 1-A: Agr., Environment* (No. E-14).

Anderson, K. and Gaston, K.J., (2013). Lightweight unmanned aerial vehicles will revolutionize spatial ecology. *Frontiers in Ecology and the Environment*, *11*(3), pp.138-146.

Bulgarelli, B., Kiselev, V. and Zibordi, G., (2014). Simulation and analysis of adjacency effects in coastal waters: a case study. *Applied optics*, *53*(8), pp.1523-1545.

Doney, S.C., Ruckelshaus, M., Emmett Duffy, J., Barry, J.P., Chan, F., English, C.A., Galindo, H.M., Grebmeier, J.M., Hollowed, A.B., Knowlton, N. and Polovina, J., (2012). Climate change impacts on marine ecosystems. *Annual review of marine science*, *4*, pp.11-37.

Feng, L. and Hu, C., (2017). Land adjacency effects on MODIS Aqua top-of-atmosphere radiance in the shortwave infrared: Statistical assessment and correction. *Journal of Geophysical Research: Oceans*, *122*(6), pp.4802-4818.

Fleischer, P., Bowker, D.E., Witte, W.G., Gosink, T.A., Hanna, W.J. and Ludwick, J.C., (1976). *Correlation of chlorophyll, suspended matter, and related parameters of waters in the lower Chesapeake Bay area to LANDSAT-1 imagery* (No. E76-10497).

Franz, B. A., Bailey, S. W., Kuring, N., & Werdell, P. J. (2015). Ocean color measurements with the Operational Land Imager on Landsat-8: implementation and evaluation in SeaDAS. *Journal of Applied Remote Sensing*, *9*(1), 096070-096070.

Gilerson, A., Malinowski, M., Herrera, E., Tomlinson, M., Stumpf, R. and Ondrusek, M., (2021). Estimation of chlorophyll-*a* concentration in complex coastal waters from satellite imagery. In *Ocean Sensing and Monitoring XIII* (Vol. 11752, pp. 64-71). SPIE.

Gitelson, A.A., Schalles, J.F. and Hladik, C.M., (2007). Remote chlorophyll-*a* retrieval in turbid, productive estuaries: Chesapeake Bay case study. *Remote Sensing of Environment*, *109*(4), pp.464-472.

Harding, L.W., Jr., E.C. Itsweire and W.E. Esaias. (1992). Determination of phytoplankton chlorophyll concentrations in the Chesapeake Bay with aircraft remote sensing. *Rem. Sens. Environ.,* 40: 79-100.

Harding Jr, L.W., Magnuson, A. and Mallonee, M.E., (2005). SeaWiFS retrievals of chlorophyll in Chesapeake Bay and the mid-Atlantic bight. *Estuarine, Coastal and Shelf Science*, *62*(1-2), pp.75-94.

IOCCG (2006). Remote Sensing of Inherent Optical Properties: Fundamentals, Tests of Algorithms, and Applications. Lee, Z.-P. (ed.), Reports of the International Ocean-Colour Coordinating Group, No. 5, IOCCG, Dartmouth, Canada

IOCCG (2008). The Societal Benefits of Ocean-Colour Technology, ser. *Reports of the International Ocean Colour Coordinating Group. Dartmouth, Canada: IOCCG*, *7*.

IOCCG (2010). Atmospheric Correction for Remotely-Sensed Ocean-Colour Products. Wang, M. (ed.), Reports of the International Ocean-Colour Coordinating Group, No. 10, IOCCG, Dartmouth, Canada.

IOCCG (2020). Synergy between Ocean Colour and Biogeochemical/Ecosystem Models. Dutkiewicz, S. (ed.), IOCCG Report Series, No. 19, International Ocean Colour Coordinating Group, Dartmouth, Canada.

Johnston, D.W., (2019). Unoccupied aircraft systems in marine science and conservation. *Annual review of marine science*, *11*, pp.439-463.

Kirk, J.T., 1994. *Light and photosynthesis in aquatic ecosystems*. Cambridge university press.

Le, C., Hu, C., Cannizzaro, J. and Duan, H., (2013). Long-term distribution patterns of remotely sensed water quality parameters in Chesapeake Bay. *Estuarine, Coastal and Shelf Science*, *128*, pp.93-103.

Magnuson, A., Harding Jr, L.W., Mallonee, M.E. and Adolf, J.E., (2004). Bio-optical model for Chesapeake Bay and the middle Atlantic bight. *Estuarine, Coastal and Shelf Science*, *61*(3), pp.403-424.

Miller, W.D. and L.W. Harding, Jr. (2007). Climate forcing of the spring bloom in Chesapeake Bay. *Mar. Ecol. Prog. Ser.,* 331: 11-22.

Munday, John C. Jr. and Fedosh, Michael S., (1981). Chesapeake Bay Plume Dynamics from LANDSAT VIMS Books and Book Chapters. 130. https://scholarworks.wm.edu/vimsbooks/130

Ondrusek, M., Stengel, E., Kinkade, C.S., Vogel, R.L., Keegstra, P., Hunter, C. and Kim, C., (2012). The development of a new optical total suspended matter algorithm for the Chesapeake Bay. *Remote Sensing of Environment*, *119*, pp.243-254.

Orth, R.J., Dennison, W.C., Lefcheck, J.S., Gurbisz, C., Hannam, M., Keisman, J., Landry, J.B., Moore, K.A., Murphy, R.R., Patrick, C.J. and Testa, J., (2017). Submersed aquatic vegetation in Chesapeake Bay: sentinel species in a changing world. *Bioscience*, *67*(8), pp.698-712.

Pahlevan, N., Smith, B., Alikas, K., Anstee, J., Barbosa, C., Binding, C., Bresciani, M., Cremella, B., Giardino, C., Gurlin, D. and Fernandez, V., (2022). Simultaneous retrieval of selected optical water quality indicators from Landsat-8, Sentinel-2, and Sentinel-3. *Remote Sensing of Environment*, *270*, p.112860.

Ridge, J.T. and Johnston, D.W., (2020). Unoccupied aircraft systems (UAS) for marine ecosystem restoration. *Frontiers in Marine Science*, *7*, p.438.

Roy, D.P., Wulder, M.A., Loveland, T.R., Woodcock, C.E., Allen, R.G., Anderson, M.C., Helder, D., Irons, J.R., Johnson, D.M., Kennedy, R. and Scambos, T.A., (2014). Landsat-8: Science and product vision for terrestrial global change research. *Remote sensing of Environment*, *145*, pp.154-172.

Shelley, P.E., (1976). *Sediment measurement in estuarine and coastal areas* (No. TR-7115-001). NASA.

Son, S. and Wang, M., (2012). Water properties in Chesapeake Bay from MODIS-Aqua measurements. *Remote Sensing of Environment*, *123*, pp.163-174.

Stumpf, R.P., (1988). Sediment transport in Chesapeake Bay during floods: analysis using satellite and surface observations. *Journal of Coastal Research*, pp.1-15.

Tomlinson, M.C., Stumpf, R.P. and Vogel, R.L., (2019). Approximation of diffuse attenuation, Kd, for MODIS high-resolution bands. *Remote Sensing Letters*, *10*(2), pp.178-185.

Tzortziou, M., Subramaniam, A., Herman, J.R., Gallegos, C.L., Neale, P.J. and Harding Jr, L.W., (2007). Remote sensing reflectance and inherent optical properties in the mid Chesapeake Bay. *Estuarine, Coastal and Shelf Science*, *72*(1-2), pp.16-32.

Vermote, E.F. and Kotchenova, S., (2008). Atmospheric correction for the monitoring of land surfaces. *Journal of Geophysical Research: Atmospheres*, *113*(D23).

Werdell, P.J., Bailey, S.W., Franz, B.A., Harding Jr, L.W., Feldman, G.C. and McClain, C.R., (2009). Regional and seasonal variability of chlorophyll-a in Chesapeake Bay as observed by SeaWiFS and MODIS-Aqua. *Remote Sensing of Environment*, *113*(6), pp.1319-1330.

Wolny, J. L., Tomlinson, M. C., Schollaert Uz, S., Egerton, T. A., McKay, J. R., Meredith, A., Reece, K.S., Scott, G.P., & Stumpf, R. P. 2020. Current and future remote sensing of harmful algal blooms in the Chesapeake Bay to support the shellfish industry. *Frontiers in Marine Science*, *7*, 337.

Zheng, G. and DiGiacomo, P.M., (2017). Remote sensing of chlorophyll-a in coastal waters based on the light absorption coefficient of phytoplankton. *Remote Sensing of Environment*, *201*, pp.331-341.

Chapter 2: Evaluating Atmospheric Correction Algorithms Applied to OLCI Sentinel-3 Data of Chesapeake Bay Waters

Introduction

At the top of the atmosphere, ocean observing satellites measure electromagnetic radiation reflected from the surface layers of water. Embedded in these measurements is water leaving radiance (L_w) which is regulated by the in-water interactions of incident irradiance (E_d) and the concentration and type of optically active constituents in the water (IOCCG, 2006). Remote sensing reflectance ($R_{rs} = L_w/E_d$) can therefore be used to derive optical properties of the water that include phytoplankton pigment concentrations (O'Reilly et al., 1998), dissolved organic matter (Mannino et al., 2008), water clarity (Lee et al., 2005), and total suspended matter (Nechad et al., 2010; Ondrusek et al., 2012). These data can significantly enhance water quality monitoring and provide a better understanding of global and local environmental processes in a changing climate.

Since the radiation source (i.e., the sun) and the satellite detector are both 'remote', the radiance incident on a given sensor is affected not only by the reflection of light from within and off the surface of water, but also by interactions with the intervening atmosphere. It is widely recognized that 90—99% of the total signal measured by a satellite sensor is due to confounding influences from the atmosphere and varying water surface facets (IOCCG, 2010). Accurately removing these influences, termed atmospheric correction, enables the science of aquatic remote sensing.

Top-of-atmosphere radiance (L_t) can be partitioned into distinct contributions from the atmosphere and ocean surface as shown in Equation (1). $L_r(\lambda)$ is radiance due to Rayleigh scattering by non-absorbing air molecules, $L_a(\lambda)$ is radiance due scattering by aerosols, $L_{ra}(\lambda)$ is the multiple interaction term between scattering by air molecules and aerosols, $L_{wc}(\lambda)$ and $L_g(\lambda)$ are radiance due to whitecaps and specular reflection of sunlight off the water surface (e.g., glint), respectively, $t(\lambda)$ and $t_0(\lambda)$ are diffuse transmittances of the atmosphere from the sun to the surface and from the surface to the sensor, respectively, $T(\lambda)$ is the direct transmittance from the surface to the sensor, and $\cos\theta_0[L_w(\lambda)]_N$ is the normalized water leaving radiance where θ_0 is the solar zenith angle. Note that path geometry (i.e., solar and viewing zenith and azimuth angles, θ, ϕ) is not shown for brevity, and λ represents wavelength.

$$L_t(\lambda) = L_r(\lambda) + L_a(\lambda) + L_{ra}(\lambda) + tL_{wc}(\lambda) + T(\lambda)L_g(\lambda) +$$
$$\mathrm{t}(\lambda)t_0(\lambda)\cos\theta_0[L_w(\lambda)]_N \quad (1)$$

Atmospheric correction schemes aim to solve a set of deterministic radiative transfer models that resolve the individual terms in Equation 1 from $L_t(\lambda)$ measurements (IOCCG, 2010; Mobley et al., 2016). Most of the partitioned radiances can be estimated with high confidence using standard radiative transfer theory and incorporating ancillary data (i.e., ozone, sea surface wind speed, atmospheric pressure measurements) (Mobley et al., 2016). In coastal waters, the main challenge is determining the aerosol contribution ($L_a + L_{ra}$) since aerosol concentrations, chemical composition, and morphology can be highly variable (IOCCG, 2010; Ahmad et al., 2010; Mobley et al., 2016). Heritage atmospheric correction schemes use aerosol models (Shettle & Fenn, 1979; Ahmad et al., 2010) to compute $L_a + L_{ra}$ as a function of the aerosol optical thickness for a variety of sensor-sun geometries (IOCCG, 2010). This process relies on the "black pixel assumption" which assumes L_w in the near-infrared (NIR) is negligible and entirely caused by atmospheric effects (Gordan & Wang, 1994; Siegel et al., 2000). Spectral variation at two NIR wavebands is used to aid in the selection of an aerosol model and computed radiances are extrapolated through the visible bands (Gordon, 1978; Gordon & Wang, 1994; Antoine & Morel, 1999; Mobley et al., 2016; IOCCG, 2010).

To date, ocean color remote sensing has been successfully exploited in open ocean waters whose optical properties are primarily determined by phytoplankton and co-varying inherent optical properties. Atmospheric correction algorithms based on the heritage Gordon (1978) approach have been used to operationally process data from most ocean color satellite sensors; however, the application of such methods has been inherently more challenging in coastal waters. While the assumption of zero NIR L_w is typically valid in the open ocean where NIR absorption by water dwarfs particle scattering, it is often invalid in productive and turbid waters where concentrations of light-scattering particles are high (Siegel et al., 2000; IOCCG, 2000; Ruddick et al., 2006). Thus, the process of extrapolating aerosol path radiance into visible bands can result in an overestimation of aerosol scattering, negative reflectance at visible wavelengths, and ultimately erroneous satellite-derived water quality products (IOCCG, 2000; Mouw et al., 2015; Zheng & DiGiacomo, 2017).

Modifications to the heritage atmospheric correction scheme have been developed to improve atmospheric correction in coastal waters. The utilization of longer wavelengths (i.e., shortwave infrared, SWIR) when present on sensors can be used to estimate and remove atmospheric signals since the absorption of water is greater and water-leaving radiance is smaller in that region (Wang & Shi, 2007; Wang, 2007; Shi & Wang, 2009; Werdell et al., 2010; Vanhellemont & Ruddick, 2015; Ibrahim et al., 2019). Other modifications to the heritage correction scheme include applying a spatially constant aerosol type, utilizing water reflectance in the red and NIR (Ruddick et al., 2000; Goyens et al., 2013), using a bio-optical model to conduct an iterative estimation of non-zero NIR (Stumpf et al., 2003; Bailey et al., 2010), and exploiting differences in the spectral shape of aerosols and water reflectance (Lavender, 2005). Alternative deterministic and statistical approaches aimed at inverting L_t in a single step have also been proposed such as spectral optimization and neural network techniques (Schroeder et al., 2007; Doerffer & Schiller, 2007; Kuchinke et al., 2009; Steinmetz et al., 2011; Hieronymi et al., 2017). Some of these approaches are developed, trained, and validated with regional datasets that may perform erroneously in other coastal water bodies.

In this study, the performance of four freely available atmospheric correction processors applied to data collected from 2016 to 2020 by the Ocean and Land Colour Instrument (OLCI) onboard the Sentinel-3 satellites are evaluated against *in situ* radiometric measurements collected in the optically complex waters of the Chesapeake Bay. OLCI currently contains the most spectral bands and the highest spatial resolution (300 m) of the current suite of missions designed for ocean color. The constellation of OLCI also provides nearly daily coverage, making it an ideal sensor to study complex, small-scale dynamic processes typical of a coastal water body.

Several studies have assessed the efficacy of various atmospheric correction schemes applied to OLCI data collected in regional coastal waters by evaluating satellite retrievals against *in situ* R_{rs} measurements (Mograne et al., 2019; Gossn et al., 2019; Kyryliuk & Kratzer, 2019; Pereira-Sandoval et al., 2019; Renosh et al., 2020; Vanhellemont & Ruddick, 2021; Giannini et al., 2021). Results from these studies demonstrate that the selection of a best performing atmospheric correction algorithm varies due to the variability of optical properties found in coastal waters globally. For example, Mograne et al. (2019) and Vanhellemont & Ruddick (2021) applied the same atmospheric correction algorithms in French and Belgian coastal waters, respectively, and obtained contrasting performance results likely due to differences in turbidity concentrations. The present study contributes to this body of literature by assessing atmospheric correction algorithms to the coastal and estuarine waters of Chesapeake Bay.

Methods

Study area

The Chesapeake Bay is a highly productive estuary located in the Mid-Atlantic region of the United States. The Bay is a shallow, partially mixed, temperate estuary spanning ~300 km in length and receives inputs of freshwater from an extensive 165,000 km^2 watershed (Kemp et al., 2005). The Bay follows a salinity gradient, with the Upper and Lower Bay largely oligohaline and polyhaline, respectively. The winter-spring freshet from the Susquehanna River typically regulates the spring phytoplankton bloom (Miller et al., 2006), while other non-algal constituents correlate with river discharge (Harding et al., 2005). Water properties can also be influenced by estuarine circulation, frontal features, sediment resuspension, and tidal cycles. Consequently, optical properties of the Chesapeake Bay waters are strongly influenced by varying concentrations of phytoplankton, total suspended matter (TSM), and chromophoric dissolved organic matter (CDOM) (Tzortziou et al., 2007). R_{rs} at 865 nm (NIR wavelength) demonstrates the wide optical variability and can be used to infer a TSM gradient in tributaries (Figure 1).

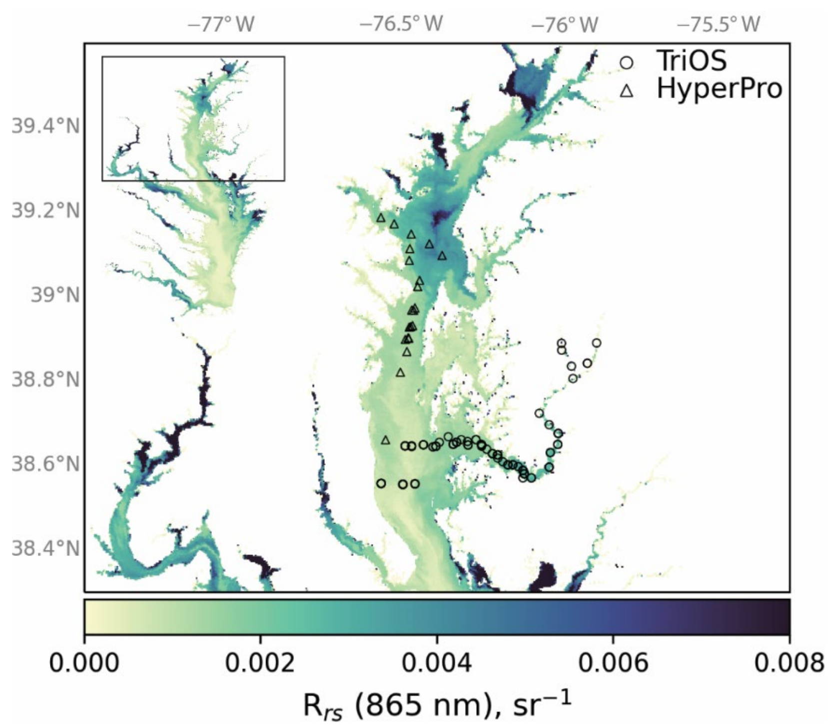


Figure 1. Study region of Chesapeake Bay, United States depicting R_{rs} at 865 nm (NIR wavelength) captured on March 23 2019 processed by the C2RCC atmospheric correction algorithm. *In situ* observation locations as collected by hyperspectral TriOS (circles) and HyperPro (triangles) radiometers.

In situ radiometry

Hyperspectral *in situ* radiometry was collected at a total of 172 stations in the Upper and Middle Bay and Choptank River, a tributary located on the eastern shore. Data was collected in different seasons and conditions between 2016 and 2020 using two methods (Figure 1). One method used a set of intercalibrated hyperspectral radiometers (TriOS RAMSES, Rastete DE) deployed on a mini-catamaran float platform (Froidefond & Ouillon, 2005) with a skylight-blocked approach (Ahn et al., 1999; Lee et al., 2013; Lee et al., 2019). The float held a downwelling irradiance sensor (E_d) and an inverted radiance sensor equipped with a plastic cone that extended fully beneath the air-water interface while the sensor fore-optics remained in the air, effectively blocking the reflected skylight contribution (L_w). The TriOS radiometers collect 256 wavelength bands at 3.3 nm intervals within the calibrated 320-950 nm range. This method collected L_w and E_d at 151 locations; *in situ* R_{rs} was calculated as L_w / E_d. The median R_{rs} was calculated from replicate measurements and any spectrum deviating more than ±10% was excluded in order to eliminate outliers due to variable measurement and illumination conditions.

The other method involved collecting underwater radiometry based on vertical profiling and propagating measurements of upwelling radiance across the water-air interface. Hyperspectral measurements were collected using a Satlantic HyperPro II (Satlantic is now Sea-Bird) free-falling profiler equipped with hyperspectral radiometers (Satlantic/Sea-Bird HyperOCRs) at 21 locations located in the Upper Bay (Ondrusek et al., 2012; 2014). The profiler was equipped with a downward directed radiance sensor to measure upwelling radiance (L_u) and an inverted upward irradiance sensor (E_d). An upward directed irradiance sensor was mounted on a gimbal at the top of the vessel to measure reference water surface irradiance (E_s) and avoid any structure shadow. The HyperOCR radiometers collect 256 wavelength bands at 3.3 nm intervals within the calibrated 350-800 nm range. The mean value of replicate casts was taken at each station. L_w was calculated just above the surface using Equation 2 where $\rho(\lambda,\theta)$ is the Fresnel reflectance index of seawater and was set to 0.021 and n^2_w is the Fresnel refractive index of seawater is set to 1.345.

$$L_w(0^+, \lambda) = L_u(0, \lambda) * [(1-\rho(\lambda,\theta))/n^2_w(\lambda)] \qquad (2)$$

All *in situ* hyperspectral measurements were interpolated at 1 nm intervals and spectrally weighted to match either the Sentinel-3A and 3B OLCI spectral response functions[1], depending on satellite matchup (see section 2.5). Due to differing approaches currently considered for the bidirectional reflectance distribution reflectance (BRDF) in coastal waters, *in situ* R_{rs} data was not BRDF corrected (see

[1] OLCI Spectral Response Function Data. Accessed online 13 March 2022: https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-3-olci/olci-instrument/spectral-response-function-data

section 2.4). Solar zenith angles were on average 33.7° (±14.3°) during *in situ* data collection.

Satellite data collection

The OLCI sensor onboard the European Commission Copernicus programmes Sentinel-3A (launched in February 2016) and Sentinel-3B (launched in April 2018) was built based upon the heritage of the Medium Resolution Imaging Spectrophotometer (MERIS). OLCI is a multispectral push-broom imaging spectrometer with five camera modules that collect top-of-atmosphere (TOA) radiance in 21 wavelengths in the 400 to 1020 nm range at a 300 m spatial resolution (EUMETSAT, 2018). Both satellites sun-synchronously orbit at a mean altitude of 815 km, providing global coverage approximately every two days (EUMETSAT, 2018). OLCI provides significant improvements to MERIS including an increase in the number of spectral bands, improved signal-to-noise, sun-glint mitigation through camera tilt, and higher spatial, temporal, and radiometric resolution (EUMETSAT, 2018).

OLCI imagery overlapping the Chesapeake Bay region (36.8, –77.35, 39.65, –75.6) from April 2016 to April 2021 were obtained from various data portals (Copernicus Online Data Access REProcessed, EUMETSAT Data Center Long-Term Archive, and Copernicus Online Data Access). A total of 2,152 scenes of TOA, full resolution (EFR), non-time critical Level-1 data and 2,119 scenes of full resolution (WFR), non-time critical Level-2 data were downloaded and stored locally. Full

Level-1 and Level-2 OLCI scenes were clipped to the Chesapeake Bay region using the Sentinel Application Platform's (SNAP) Graph Processing Tools (GPT, 'Subset').

Atmospheric correction algorithms

Four atmospheric correction algorithms were evaluated against *in situ* measurements: The Baseline Atmospheric Correction (BAC) v7.0, the Case 2 Regional Coast Color (C2RCC) v7.0, the Polynomial-based algorithm applied to MERIS (POLYMER) v4.12, and NASA's standard Level-2 generator atmospheric correction algorithm (L2gen) v7.5.3 (Figure 2). Downloaded Level-2 data were pre-processed with BAC, and Level-1 data were locally processed with the other atmospheric correction algorithms. Additional information on processing parameters for each algorithm can be found in Figure 2 and Tables S1-S3. Due to differing approaches currently considered for the BRDF in coastal waters, and following methods in similar optical complex R_{rs} matchup studies (Vanhellemont & Ruddick, 2021; Giannini et al., 2021), BRDF correction was excluded from all atmospheric correction schemes.

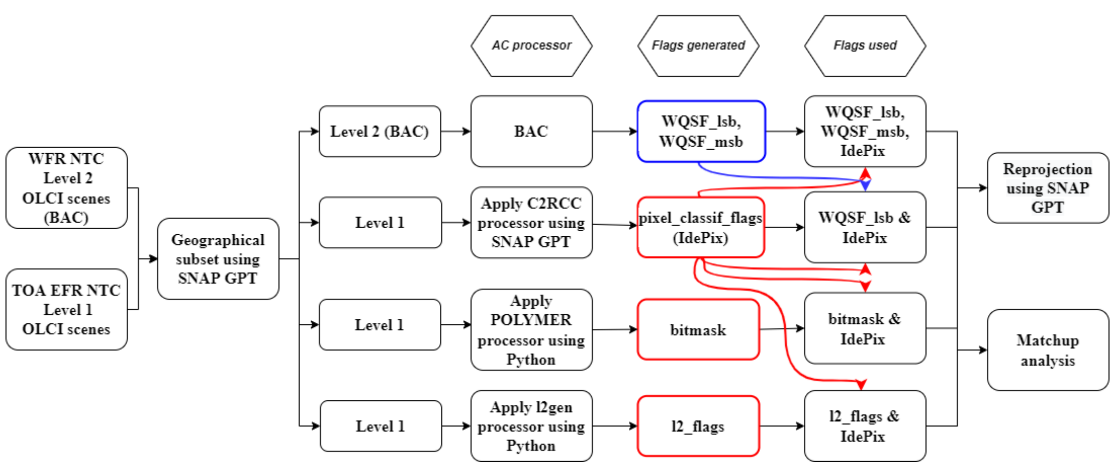


Figure 2. Processing and flagging workflow of four different atmospheric correction processors.

The BAC algorithm is based on the basic principles for the MERIS atmospheric correction scheme which was designed for open ocean waters and uses a multiple scattering algorithm to assess the aerosol contribution (Antoine & Morel, 1999; IOCCG, 2010). BAC integrates a Bright Pixel Atmospheric Correction (BPAC) to account for when NIR water leaving radiance is not negligible, such as in high scattering waters (Moore et al., 1999; Lavender et al., 2005; Moore et al., 2017). A detected NIR signal partitions the TOA reflectance into components due to aerosols and TSM using a coupled hydrological and atmospheric model (Moore et al., 2017). Validation of this approach was based on MERIS match-ups with *in situ* data provided by the MERIS Matchup *In situ* Database (MERMAID) which includes compiled global *in situ* radiometry data (Barker et al., 2008; Moore et al., 2017).

The C2RCC algorithm, originally developed by Doerffer and Schiller (2007) for MERIS, is based on an artificial neural network (ANN) inversion model to derive water leaving radiance and in-water optical properties. The ANN is trained on a large database of simulated angle-dependent water leaving reflectances and TOA radiances generated by the radiative transfer model HydroLight (Mobley, 1994). TOA radiances at fifteen wavelengths are corrected for absorbing gasses and the smile effect and used together with ancillary data (i.e., wind speed, salinity, temperature, viewing geometry) to perform an inversion of water-leaving radiance and atmospheric parameters (Doerffer & Schiller, 2007; Brockmann et al., 2016). C2RCC v7.0 was applied locally to Level-1 OLCI data using SNAP GPT. C2RCC v7.0 is updated with new default neural nets (i.e., NNv2) that fully supersedes the older neural net[2]. The most updated system vicarious calibration (SVC) gains were applied (Table S2) and the derived retrievals were not corrected for BRDF effects.

The POLYMER algorithm developed by HYGEOS (Lille, France) was also initially developed for MERIS and consists of a full-spectrum coupled spectral matching algorithm (Steinmetz et al., 2011). The algorithm relies on a bio-optical reflectance model of Park & Ruddick (2005) and can be modified with known chlorophyll *a* concentrations and particle backscattering coefficients to represent various oceanic and coastal waters (Soppa et al., 2021). A polynomial expression is used to represent the coupled atmosphere and water-leaving reflectance and an iterative process is used to optimize parameters to obtain the best spectral fit of water leaving reflectance (Steinmetz et al., 2011; Steinmetz & Ramon, 2018; Soppa et al., 2021). This algorithm determines aerosol contributions from a linear combination of reflectance terms, rather than from a specific aerosol model, and can characterize complex atmospheric and surface effects such as residual sun glint. POLYMER was trained on *in situ* data collected from the North Sea (Kratzer & Plowey, 2020) and has been validated using global AERONET-OC sites (Steinmetz & Ramon, 2018). In this study, POLYMER was locally processed using Python v3.7. All POLYMER defaults

[2] Sentinel-3 Product Notice. Accessed 13 March 2022, https://www.eumetsat.int/media/48139

were kept, including the hard-coded SVC gains (Table S2); however, BRDF correction was disabled (Table S1).

L2gen is the standard atmospheric correction algorithm that the NASA Ocean Biology Processing Group (OBPG) employs which builds on several decades of atmospheric correction research (Gordon & Wang, 1994; Stumpf et a., 2003; Bailey et al., 2010; Ahmad et al., 2010). The current version of L2gen consists of using precomputed radiative transfer simulations, models, and ancillary information in order to correct TOA radiance for atmospheric terms such as gaseous transmittance, white caps, Rayleigh scattering, and sun glint (Mobley et al., 2016). The aerosol contribution is estimated using two NIR wavelengths where water leaving radiance can be accurately extrapolated to the visible wavelengths (Bailey et al., 2010). For waters with non-negligible water-leaving reflectance in the NIR, an optical model is used to estimate NIR, and an iterative bio-optical modeling approach is used to extrapolate aerosol contribution to the visible wavelengths (Bailey et al., 2010). In this study, L2gen was locally processed using Python v3.7 with the L2gen aerosol mode option selected as multi-scattering with 2-band model selection and iterative NIR correction using OLCI wavelengths 865 nm and 1012 nm (Table S1). SVC gains were applied according to OBPG recommendations and BRDF effects were not corrected for (Tables S1 and S2).

All atmospherically corrected products were flagged using recommended quality flags for each atmospheric correction processor (Figure 2, Table S3). The Identification of Pixels (IdePix) pixel identification tool was run during the processing of C2RCC, and associated flags were applied to all data for comprehensive image masking (Figure 2, Table S3). IdePix is a SNAP operator that calculates pixel-by-pixel cloud probability; however, other pixel types such as cloud shadow, coastline, or land can also be masked. IdePix is recognized to reliably mask cloudy pixels or pixels of mixed surface types and was applied to all atmospheric correction processors to enhance flagging results, similar to other atmospheric correction evaluations (Sathyendranath et al., 2019; Kratzer & Plowey, 2021). The WQSF_lsb data contains quality flags that are relevant to both the BAC and C2RCC processing; therefore, relevant WQSF_lsb quality flags were applied to each, in line with EUMETSAT match-up protocol guidance. In addition to IdePix, all the recommended quality flags were used for POLYMER and L2gen.

Match-up analysis

Satellite-derived R_{rs} measurements were co-located in time and space in comparison with *in situ* measurements following EUMETSAT recommendations (EUMETSAT, 2021). Satellite scenes corresponding to the same date and within 3 hours of an *in situ* measurement were obtained. The nearest satellite pixel located ≤ 250 m from an *in situ* measurement location was identified and R_{rs} was extracted from a 3 x 3 full resolution pixel window (900 m^2) centered on the nearest pixel location. R_{rs} values from these matchups were included if 50% + 1 pixels in the 3 x 3 pixel window (5 out of 9) were valid (not flagged). Pixel outliers were identified as mean satellite R_{rs} measurements greater or less than 1.5 times the standard deviation and were removed from the analysis. To ensure homogeneity, matchups were also discarded if the coefficient of variation at 560 nm was greater than 20%. Median and standard deviation R_{rs} values were extracted and compared to *in situ* values using various matchup statistics following EUMETSAT recommendations (EUMETSAT, 2021). Statistical metrics included linear regression statistics (slope, intercept, coefficient of determination (r^2)) as well as root mean square error (RMSE), mean deviation (MD) and mean percentage deviation (MPD) to investigate prediction errors, bias and dispersion, respectively.

$$MD = \frac{\sum_{i=1}^{n}(Rrs_{Sat} - Rrs_{in\ situ})}{n} \qquad (3)$$

$$MPD = \frac{\sum_{i=1}^{n} \frac{100(Rrs_{Sat} - Rrs_{in\ situ})}{Rrs_{in\ situ}}}{n} \qquad (4)$$

Performance metrics between *in situ* and satellite-derived R_{rs} for each atmospheric correction algorithm were also analyzed using Taylor Diagrams (Taylor, 2001). These polar coordinate diagrams provide a compact summary of how three complementary model performance statistics vary simultaneously (Jolliff et al., 2008). The correlation coefficient, standard deviation, and RMSE are related to one another through the Law of Cosines and can be plotted together on one 2D graph (Taylor, 2001). The magnitude of the variability is indicated by computed standard deviation and illustrated as the radial distance from the origin of the plot with points farthest from the origin containing higher variability. Correlation coefficients are shown on the arc of the coordinate plot with points closest to the *x*-axis containing the highest correlation. RMSE is indicated by the concentric dashed lines emanating from the origin where points farthest from the observed value have high RMSE.

OLCI composites

To investigate occurrence and spatial variability of quality flags, composites of clear (unflagged) pixels were created for each atmospheric correction. The number of unflagged pixels in all calendar year 2020 data were summed and mapped to illustrate the spatial variability of 'clear day' spatial coverage across Chesapeake Bay. Composites were produced by merging 2020 OLCI scenes of the same day processed by each atmospheric correction algorithm (n = 336, 307, 353, 355 for BAC, C2RCC, POLYMER, L2gen, respectively). Data were reprojected to the World Geodetic System 1984 coordinate system (SNAP GPT 'Reproject') and the quality flags used in the matchup analysis were applied (Figure 2, Table S3). The 'xarray' Python module was used to create climatology datasets. Seasonal R_{rs} climatologies for individual wavelengths were produced by averaging pixel values across each season (Dec-Feb, Mar-May, Jun-Aug, Sep-Nov).

Results

In situ and satellite-derived radiometry

In general, hyperspectral *in situ* R_{rs} spectra were representative of a eutrophic system (Figure 3, Gitelson et al., 2007; Gitelson et al., 2008; Spyrakos et al., 2018).

Due to absorption of CDOM and chlorophyll *a* in lower wavelengths, R_{rs} is low in the blue region (400–500 nm) and higher in the green region (500–600 nm). The red region (600–700 nm) decreases and contains a second minimum around 675 nm corresponding to the red chlorophyll *a* absorption maximum (Gitelson et al., 2008). The rise in R_{rs} between 690 and 715 nm is the result of high backscattering and a minimum in absorption by all optically active components besides pure water (Gitelson et al., 2008). This secondary peak can also be attributed to chlorophyll fluorescence emission (Spyrakos et al., 2018). Some R_{rs} spectra are high in the NIR (700–800 nm) due to highly scattering particulate matter (Gitelson et al., 2008).

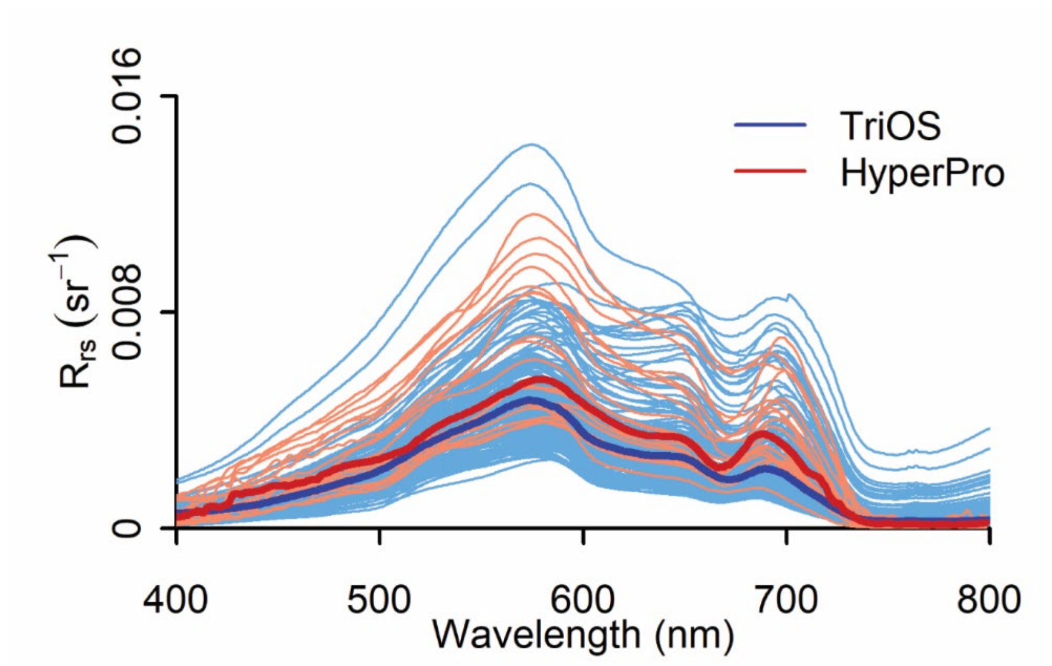


Figure 3. *In situ* R_{rs} spectra collected with above-water hyperspectral TriOS radiometers (blue) and free-falling HyperPro radiometers (red) at 172 stations in the Upper and Middle Chesapeake Bay and Choptank River tributary between April 2016 and April 2020. The bold spectra represent the median R_{rs} spectra for each *in situ* collection method.

Matchup analysis

A time difference of ≤ 3 hours with flagging applied resulted in 27 matchups for BAC, 41 for C2RCC, 48 for POLYMER, and 46 for L2gen. Comparisons of the matchup results to *in situ* R_{rs} are presented as mean spectral comparisons, scatterplots, and Taylor Diagrams in Figures 4–6 respectively. Averaged across the matchup dataset, C2RCC and L2gen data are most similar to the *in situ* R_{rs} spectra across all OLCI wavebands. Conversely, BAC and POLYMER data overestimate *in situ* R_{rs}, particularly in the green and red spectral region (Figure 4). All atmospherically corrected R_{rs} tend to agree towards the longer wavelengths (681–779 nm) with lower standard errors (Figure 4).

Figure 5 and Table 1 demonstrate the accuracy of atmospherically corrected R_{rs} is wavelength-dependent as the correlation coefficient decreases with wavelength across all AC algorithms (Figure 5, Table 1). Regression slopes are low or negative and correlations of determination are weak in the shorter wavelengths (400–443 nm, slope = –36.28 - 0.01, r^2 = 0.00–0.08) and increase in the longer wavelengths (665–779 nm, slope = 0.33–8.22, r^2 = 0.28–0.78). C2RCC and L2gen are the only processors that do not result in negative regression slopes. BAC and C2RCC contained the highest correlations of determination (BAC at 665 nm, r^2 = 0.82, C2RCC at 754 nm, r^2 = 0.72). Mean RMSE and mean deviation was highest in POLYMER and BAC across all wavelengths (POLYMR mean RMSE = 0.017, POLYMER mean MD = 0.006, BAC mean RMSE = 0.011, BAC mean MD = 0.005). C2RCC had the lowest mean RMSE and mean deviation (C2RCC mean RMSE =

0.001, mean MD = 0.001). Mean percentage deviation was overall higher for POLYMER and BAC across all wavelengths (POLYMER mean MPD = 1630, BAC mean MPD = 608) and lowest for C2RCC (mean MPD = 131).

The location of points in the Taylor Diagrams in Figure 6 demonstrate how C2RCC had the lowest standard deviation at all wavelengths (points are close to the origin), contained the highest correlation across all wavelengths (points are close to *x*-axis), and had the smallest unbiased RMSE (points are close to *in situ* value). Taylor plots portray L2gen performing second best with POLYMER and BAC performing the poorest across all bands.

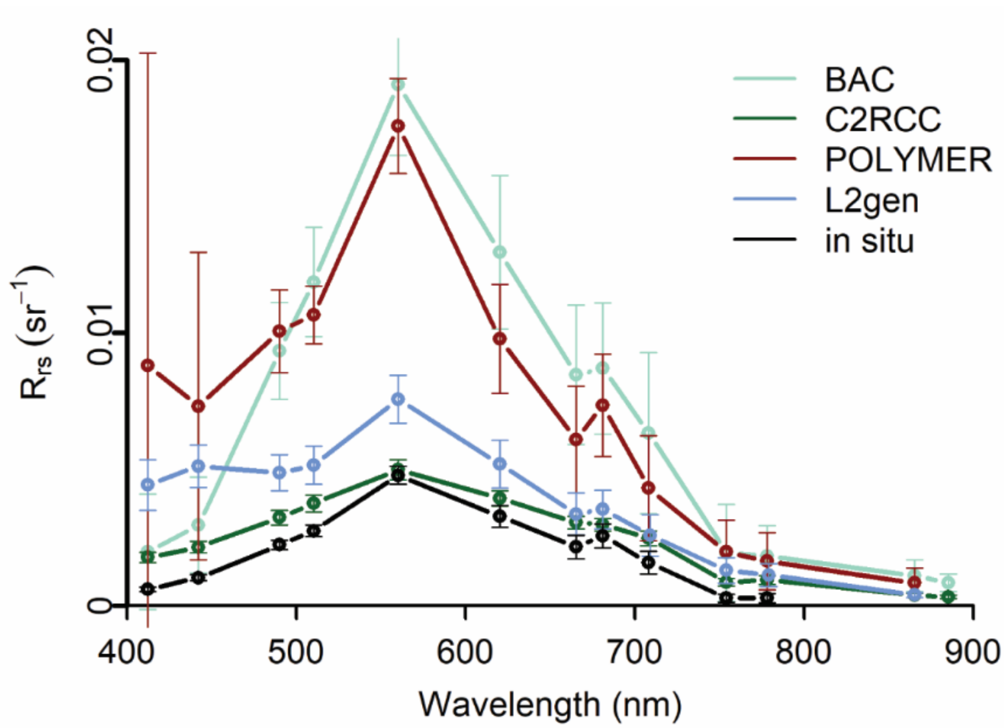


Figure 4. Median satellite-derived R_{rs} processed by each atmospheric correction algorithm and *in situ* R_{rs} spectra (black) across all matchups. Error bars represent standard error.

Table 1. Statistical matchup results for satellite-derived R_{rs} processed by each atmospheric correction algorithm against *in situ* R_{rs}.

	BAC (n = 27)						C2RCC (n = 41)					
λ (nm)	**Slope**	**Int**	**r^2**	**RMSE (sr^{-1})**	**MD (sr^{-1})**	**MPD**	**Slope**	**Int**	**r^2**	**RMSE (sr^{-1})**	**MD (sr^{-1})**	**MPD**
400	−7.97	0.009	0.07	0.009	0.002	1470	0.03	0.002	0	0.001	0.001	303
412	−5.73	0.008	0.05	0.008	0.001	1130	0.02	0.002	0	0.001	0.001	236
443	−2.8	0.009	0.03	0.008	0.002	548	0.01	0.002	0	0.001	0.001	119
490	0.47	0.009	0	0.010	0.006	397	0	0.003	0	0.002	0.001	64
510	1.85	0.008	0.04	0.013	0.008	365	0.11	0.003	0.01	0.002	0.001	47
560	4.01	0.001	0.25	0.018	0.014	333	0.3	0.004	0.08	0.002	0	24
620	5.89	−0.006	0.71	0.015	0.009	311	0.28	0.003	0.16	0.001	0	29
665	5.43	−0.003	0.82	0.012	0.006	230	0.33	0.002	0.32	0.001	0	34
681	4.93	−0.003	0.78	0.011	0.006	298	0.32	0.002	0.3	0.001	0	30
709	5.98	−0.002	0.72	0.013	0.005	396	0.48	0.002	0.47	0.001	0.001	51
754	8.22	0.001	0.5	0.007	0.002	1043	0.94	0.001	0.72	0.001	0	249
779	4.52	0.001	0.48	0.005	0.002	774	0.95	0.001	0.7	0.001	0.001	378
Mean	**2.07**	**0.003**	**0.37**	**0.011**	**0.010**	**608**	**0.31**	**0.002**	**0.23**	**0.001**	**0.001**	**130**
	POLYMER (n = 48)						**L2gen (n = 46)**					
λ (nm)	**Slope**	**Int**	**r^2**	**RMSE (sr^{-1})**	**MD (sr^{-1})**	**MPD**	**Slope**	**Int**	**r^2**	**RMSE (sr^{-1})**	**MD (sr^{-1})**	**MPD**
400	−36.28	0.054	0.04	0.060	0.01	9809	−4.26	0.008	0.11	0.006	0.005	1514
412	−30.19	0.041	0.05	0.046	0.008	4483	−3.14	0.007	0.08	0.005	0.004	790
443	−8.32	0.024	0.05	0.024	0.006	1403	−0.34	0.006	0.01	0.005	0.004	394
490	1.13	0.008	0.04	0.010	0.008	395	0.5	0.004	0.04	0.004	0.003	154
510	2.08	0.005	0.43	0.009	0.008	277	0.73	0.003	0.13	0.004	0.002	103
560	2.54	0.003	0.6	0.013	0.012	233	0.98	0.002	0.36	0.004	0.003	60
620	2.8	−0.001	0.53	0.009	0.006	192	1.25	0	0.55	0.003	0.002	55
665	2.85	−0.001	0.49	0.007	0.004	205	1.21	0.001	0.55	0.002	0.001	59
681	2.71	−0.001	0.47	0.007	0.005	194	1.08	0.001	0.51	0.002	0.001	51
709	2.98	−0.002	0.49	0.006	0.003	196	1.33	0	0.61	0.002	0.001	667
754	4.12	0.001	0.29	0.004	0.002	1122	1.84	0.001	0.35	0.002	0.001	719
779	3.65	0.001	0.35	0.004	0.001	1059	1.54	0.001	0.39	0.002	0.001	757
Mean	**−4.16**	**0.011**	**0.32**	**0.017**	**0.006**	**1631**	**0.23**	**0.003**	**0.31**	**0.003**	**0.002**	**444**

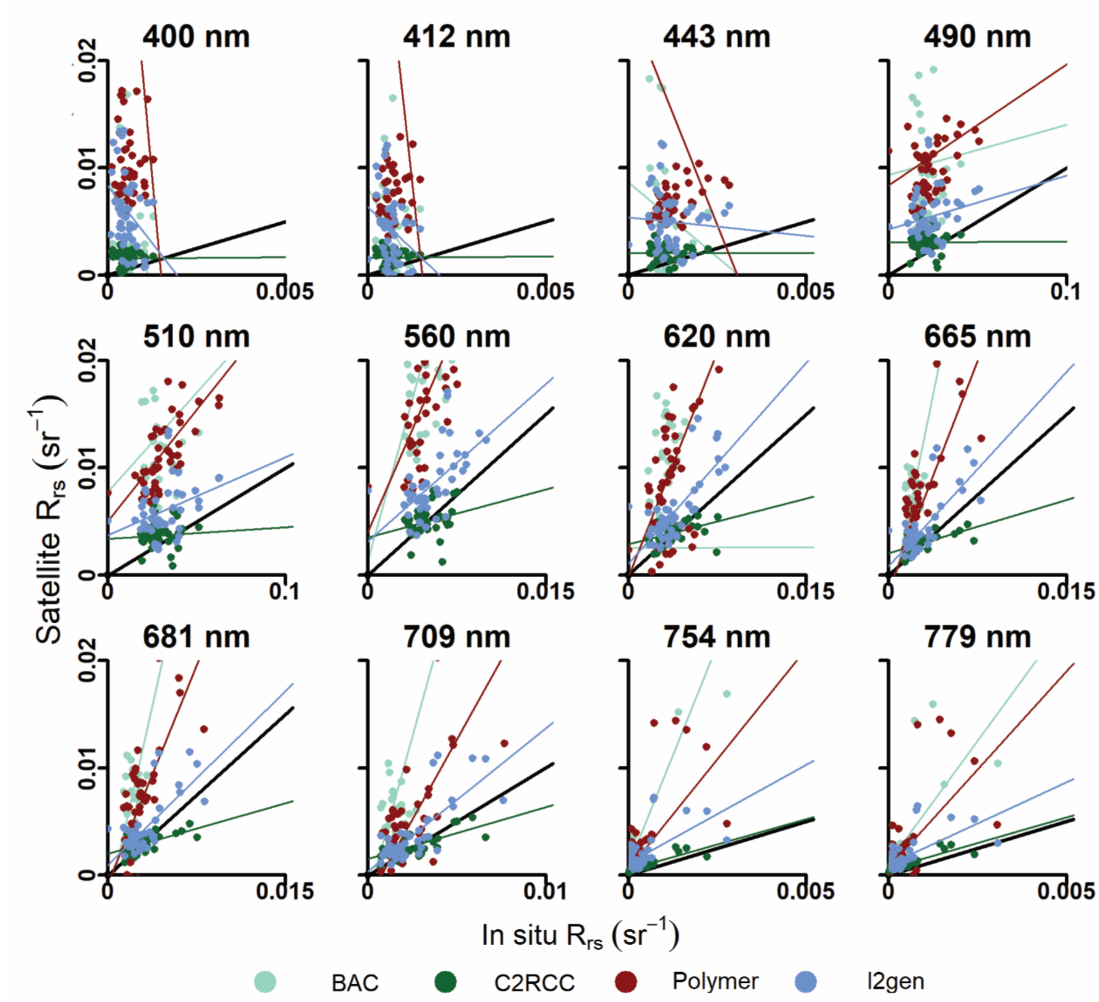


Figure 5. Scatter plots of *in situ* R_{rs} vs satellite-derived R_{rs} for OLCI wavelengths ranging from 400 to 779 nm. The black line represents the 1:1 line and colored lines are linear regression fits for each atmospheric correction processor.

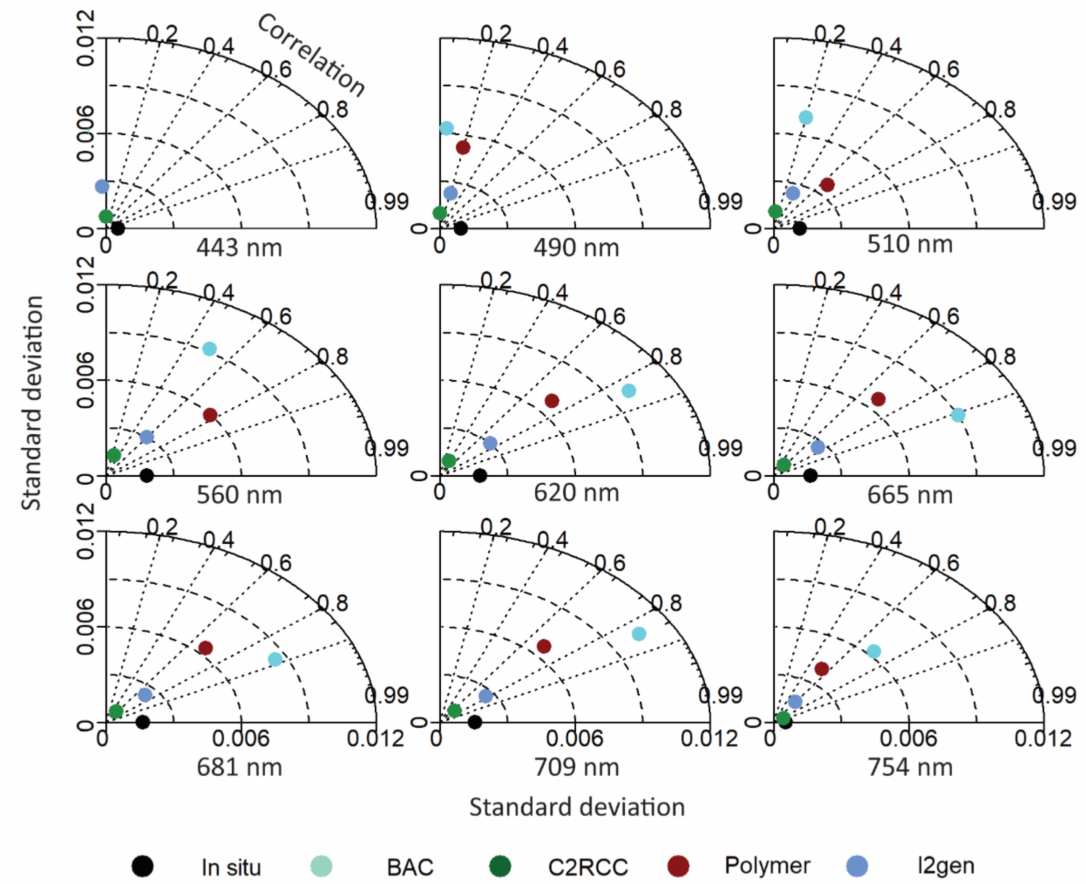


Figure 6. Taylor diagrams of atmospherically corrected R_{rs} in comparison to *in situ* R_{rs}. Each panel shows separately the results for nine OLCI wavelengths ranging from 443 to 754 nm. Standard deviation is illustrated as the radial distance from the origin, correlation coefficients are shown on the arc of the coordinate plot with points closest to the x-axis containing the highest correlation, and RMSE is indicated by the concentric dashed lines emanating from the origin where points farthest from the observed value have high RMSE error.

Pixel flagging

The spatial variability of pixel flagging for each atmospheric correction processor demonstrates that more pixels are flagged (reduction of clear day pixels) in the Upper Bay and along the coastline (Figure 7). Across all 2020 OLCI scenes of only Chesapeake Bay water pixels, IdePix flagged 46% of pixels as clouds and 59% as associated cloud pixels (i.e., cloud margin, cloud shadow, etc.), 39% as bright pixels, and 4% as land although land was previously masked out. The BAC quality flags (WQSF_lsb) flagged 57% of water pixels as cloud associated flags, 8% as high glint, and 3% as invalid or 'AC fail'. The BAC science flags (WQSF_msb) flagged 2–11% of water pixels as negative water-leaving radiance for wavebands 412–665 nm, with more flagging at the lower wavelengths. POLYMER quality flags flagged 49% of water pixels as clouds, 4% as invalid and 1% as thick aerosol. L2gen flags flagged 61% as clouds, 9% as atmospheric correction warning, 6% as high glint, and 14% as land although the land was previously masked out. The default land mask for L2gen is approximately 500 m which leads to masking out water pixels in narrow tributaries and could be rectified by passing a null landmask or defining a custom mask (Sean Bailey, personal communication).

Clear day coverage is lowest for the L2gen processing workflow and highest for the POLYMER processing workflow (Figure 7). BAC flagging produced a mean of 52 clear days with higher coverage in the Middle and Lower Bay and less in the Upper Bay (Figure 7). C2RCC flagging produced an average of 58 clear days and is distributed across the entire Bay. POLYMER flagging produced an average of 85 clear days which, like BAC, has higher coverage in the Middle and Lower Bay. L2gen flagging produced an average of 47 clear days which is more prominent in the Lower Bay (Figure 7). More pixels are flagged and removed in the Upper Bay and tributaries likely due to increased challenges in atmospheric correction and more optically complex waters. Pixels adjacent to the Bay shoreline are also increasingly flagged, likely due to the adjacency effect where nearby bright pixels are scattered into the field of view of neighboring water pixels.

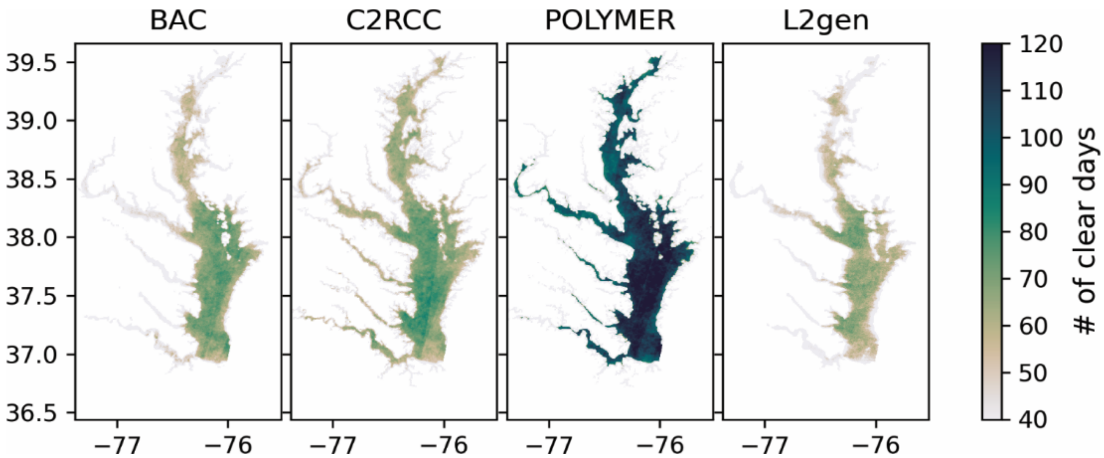


Figure 7. Number of clear (unflagged) pixels summed for 2020 for each atmospheric correction algorithm.

R_{rs} composites

R_{rs} composites of the 665 nm OLCI waveband demonstrate differences in R_{rs} across time and space for each atmospheric correction algorithm (Figure 8). In general, BAC contains the highest R_{rs} followed by POLYMER and C2RCC while L2gen produced lower overall R_{rs} (Figure 8). In all atmospheric correction algorithms, R_{rs} at 665 nm tends to be high in the Upper Bay and tributaries, especially during winter and spring owing due to elevated suspended matter concentrations.

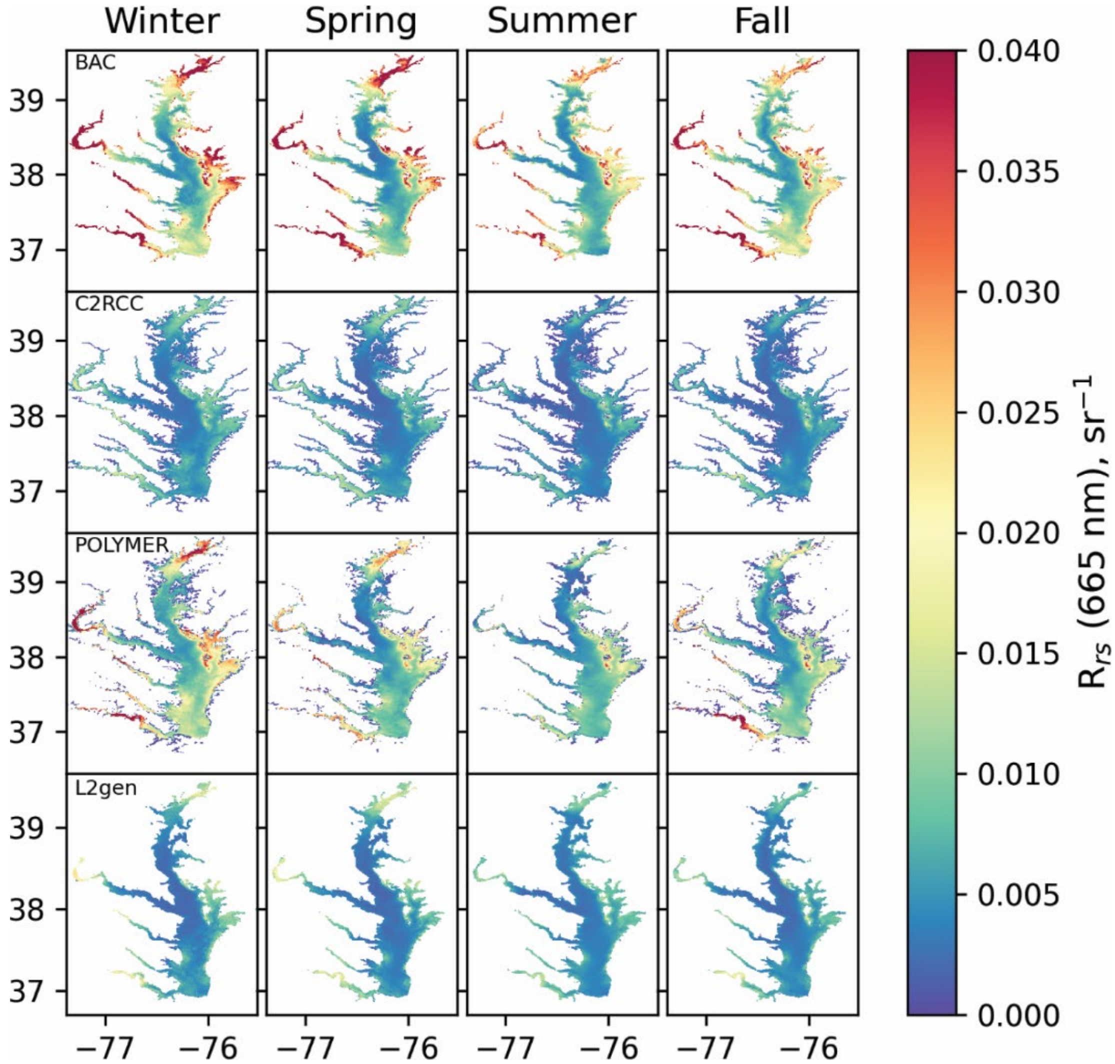


Figure 8. Seasonal climatologies of R_{rs} at 665 nm for the year 2020 derived by each atmospheric correction algorithm.

Discussion

This study is the first attempt to evaluate atmospheric correction performance applied to Sentinel-3A and 3B OLCI scenes of Chesapeake Bay waters. Overall, C2RCC performed best when compared to *in situ* data collected in the middle and upper Bay. Results have broad implications for improving satellite water quality retrievals in Chesapeake Bay, however, several aspects and caveats of this study merit additional discussion.

Performance of atmospheric correction processors in Chesapeake Bay

In general, the accuracy of R_{rs} retrievals from the atmospheric correction processors was wavelength-dependent, with performance increasing at longer wavelengths (Table 1, Figures 4 and 5). The spectral shapes of R_{rs} averaged across all valid matchup locations were similar to *in situ* spectra and typical reflectance spectra collected in coastal waters (Gitelson et al., 2008). However, the magnitude of mean R_{rs} spectra varied across atmospheric correction algorithms (Figure 4). The range of R_{rs} in the green waveband (560 nm) demonstrates high variability in scattering by inorganic and organic particles (e.g., suspended sediment and phytoplankton). All algorithms produced a distinct secondary peak at 681 nm associated with the chlorophyll *a* absorption maximum and potentially fluorescence, except for C2RCC (Figure 4). Vanhellemont & Ruddick (2021) hypothesize that the C2RCC neural net is limited by its training dataset and in high reflectance waters, resorts to using spectra containing an absorption dip at 681 nm. Some *in situ* locations may have been in areas of high reflectance, prompting the neural net to use spectra that when averaged, resulted in flattening at 681 nm. Therefore, water quality products derived using this wavelength could potentially be erroneous. The BAC and L2gen algorithms that use a two-band iterative approach likely had higher errors in the blue wavelengths associated with extrapolating the aerosol contribution from a NIR band to the visible (IOCCG, 2010; Frouin et al., 2019). Similar performance was reported in previous studies using L2gen on Chesapeake Bay scenes collected by NASA's Moderate Resolution Imaging Spectroradiometer (MODIS) (Werdell et al., 2010; Son & Wang, 2012). These authors concluded that low signal-to-noise on NIR and SWIR bands propagates into the blue bands, leading to an overestimation of aerosols and negative or low R_{rs} retrievals (Werdell et al., 2010; Son & Wang, 2012). The POLYMER algorithm, which uses a spectral optimization approach, does not use a specific aerosol model like BAC and L2gen, but instead, uses the whole spectral range from the blue to NIR to decouple the atmospheric and surface components of the TOA signal from the water leaving reflectance (Steinmetz & Ramon, 2019). This spectral matching approach is based on chlorophyll *a* concentration and backscattering and does not contain a parameter directly related to non-varying CDOM absorption. Therefore, POLYMER likely overestimated R_{rs} in the blue bands due to the exclusion of highly absorbing CDOM found in Chesapeake Bay waters.

All atmospheric correction algorithms had better performance in the longer wavelengths (560–865 nm) with C2RCC containing the lowest RMSE, MD, and MPD (Table 1, Figures 4 and 5). POLYMER retrieved a few unrealistic negative values in bands 665–865 nm, likely caused by the spectral matching approach. The applicability of the assumptions made in the in-water model used in POLYMER may not be suitable for eutrophic waters, where different IOP models produce contrasting outputs with varying phytoplankton concentrations (Lain et al., 2014). R_{rs} at 665–709 nm are wavelengths that are frequently used to derive water turbidity or total suspended matter (Nechad et al., 2010), and chlorophyll *a* in coastal waters (Gons et al., 2005; Moses et al., 2019); therefore, water quality products derived from these bands are likely to be more accurate than band-ratio algorithms incorporating blue wavelengths.

In situ measurements in this study cover a dynamic range for the upper and middle Bay as well as locations in the Choptank River (Figure 3). These measurements, conducted throughout the year, are likely to be representative of waters spanning the entire Chesapeake Bay; however, the authors acknowledge the limitation in the number of samples and the need for further radiometric validation with a larger and more extensive dataset. Future studies should consider incorporating *in situ* radiometric data in the Lower Bay and measurements collected from the recently installed radiometer included in the Ocean Color Component of the NASA Aerosol Robotic Network (AERONET-OC) located at 39.13 N, –76.32 W (Zibordi et al., 2020).

Comparison to similar studies

The results of this study can be compared to other OLCI matchup analyses in coastal waters found in the literature (Table 2, Hieronymi et al., 2017; Mograne et al., 2019; Alikas et al., 2020; Vanhellemont & Ruddick, 2021; Giannini et al., 2021). However, statistical reporting across these studies is not consistent, which makes it challenging to compare atmospheric correction performance across different study sites. EUMETSAT provides guidelines for a common matchup approach to achieve a consistent validation baseline to compare across different studies (EUMETSAT, 2021), and were used in this study. Adopting these matchup statistics will provide comparable values and allow for better investigation of the performance of atmospheric correction processors across various regional coastal water bodies.

Table 2. Summary of existing OLCI atmospheric correction evaluation literature. Table information includes the atmospheric correction algorithms used, study site, *in situ* R_{rs} matchup sample size, and performance metric(s) for four wavelengths, if available. BAC: EUMETSAT's Baseline Atmospheric correction algorithm, C2RCC: EUMETSAT's Case 2 Regional Coast Color, C2R-CCAltNets: alternative neural nets of C2RCC, POLYMER: POLYnomial based algorithm applied to MERIS, L2gen: NASA's standard atmospheric correction algorithm, ACOLITE/DSF: ACOLITE atmospheric correction with dark spectrum fitting (DSF). MD: Mean deviation, MPD: Mean Percentage Deviation.

		443			560			665			754		
	n	r^2	MD	MPD	r^2	MD	MPD	r^2	MD	MPD	r^2	MD	MPD
This study													
BAC	27	0.03	0.002	548	0.25	0.014	333	0.82	0.006	230	0.50	0.002	1043
C2RCC	41	0	0.001	119	0.08	0	24	0.32	0	34	0.72	0	249
POLYMER	48	0.05	0.006	1403	0.60	0.012	233	0.49	0.004	205	0.29	0.002	1122
L2gen	46	0.01	0.004	394	0.36	0.003	60	0.55	0.001	59	0.35	0.001	719
Mograne et al., 2019													
BAC	18	0.38 *	–	−40 *	0.95 *	–	−10 *	0.95 *	–	−40 *	–	–	–
C2RCC	16	0.63 *	–	−5 *	0.93 *	–	0 *	0.95 *	–	−20 *	–	–	–
C2R-CCAltNets	19	0.84 *	–	15 *	0.94 *	–	10 *	0.90 *	–	5 *	–	–	–
POLYMER	35	0.54 *	–	−18 *	0.86 *	–	−12 *	0.93 *	–	−18 *	–	–	–
L2gen	17	0.50 *	–	22 *	0.98 *	–	−4 *	0.98 *	–	−12 *	–	–	–
Alikas et al., 2020													
BAC	22	0.028	–	−192	0.59	–	−2	–	–	12	–	–	58
C2RCC	22	0.29	–	64	0.6	–	16	–	–	39	–	–	122
C2RCC ALTNN	22	0.003	–	50	0.37	–	7	–	–	41	–	–	72
POLYMER	22	0.82	–	6	0.96	–	14	–	–	20	–	–	55
Vanhellemont & Ruddick, 2021													
BAC	46	0.61	−0.013	–	0.86	−0.008	–	0.90	0.000	–	0.73	−0.001	–
POLYMER	27	–	−0.025 *	–	–	−0.033	–	–	−0.025	–	–	−0.015 *	–
C2RCC	27	–	−0.005 *	–	–	−0.018	–	–	−0.017	–	–	−0.008 *	–
L2gen	27	–	−0.005 *	–	–	−0.012	–	–	−0.007	–	–	0.000 *	–
ACOLITE/DSF	46	0.37	−0.005	–	0.86	−0.004	–	0.90	−0.003	–	0.67	0.001	–
Giannini et al., 2021													
C2RCC	24	0.74	–	−33.6	0.83	–	−28.8	0.76	–	−36.3	–	–	–
C2RCC altNN	24	0.36	–	−36.9	0.72	–	−18.2	0.66	–	−12	–	–	–
POLYMER	24	0.73	–	−41.5	0.83	–	−23.6	0.71	–	−28.4	–	–	–

* estimated values from spectra plots of derived statistics.

Mograne et al. (2019) conducted a R_{rs} matchup analysis focused on the coastal waters of the Eastern English Channel and French Guiana (n=18, 19, 35, 17 for BAC, C2RCC, POLYMER and L2gen, respectively). For low wavelengths (400–510 nm), BAC was always negatively biased and contained the highest RMSE while C2RCC was always positively biased and contained the lowest RMSE. For longer wavelengths (560–754 nm), C2RCC also contained the lowest RMSE while POLYMER contained high RMSE. The authors recommended the use of POLYMER or C2RCC to process OLCI images over their region of interest. An R_{rs} matchup analysis focusing on the coastal waters of British Columbia (n = 24) demonstrated that C2RCC performed best since it allowed for large variability in water-leaving radiance for regions impacted by the presence of CDOM (Giannini et al., 2021).

POLYMER showed slightly inferior radiometric performance; however, it was the most accurate when comparing Level-2 operational products (e.g., chlorophyll *a* and TSM) products to *in situ* measurements. The authors hypothesize this was due to the preservation of the spectral shape in R_{rs} compared to C2RCC (Giannini et al., 2021). Hieronymi et al. (2017) developed the OLCI Neural Network Swarm (ONNS) which consists of several specialized neural networks to derive ocean color remote sensing products in a wide range of optical properties. This study applied ONNS to C2RCC corrected OLCI data and compared derived water quality concentrations to *in situ* measurements (n = 43). Results of this study showed relatively good agreement to *in situ* data, particularly for inorganic suspended matter and CDOM (r = 0.86 and 0.73, respectively)

Results of this study also contrast with similar matchup studies focused in coastal waters (Alikas et al., 2020; Vanhellemont & Ruddick, 2021). An OLCI R_{rs} matchup in the Belgian coastal zone (n = 46) demonstrated poor performance in POLYMER and C2RCC while BAC performed relatively well (Vanhellemont & Ruddick, 2021). Vanhellemont & Ruddick (2021) hypothesize that POLYMER and C2RCC are constrained by the bio-optical models and are unable to represent higher reflectance in turbid water, resulting in an underestimation of R_{rs} at red wavelengths. These results are not consistent with the present study; however, *in situ* locations were also not located in highly turbid waters. Future work should test the accuracy of C2RCC and POLYMER in more turbid conditions in Chesapeake Bay. Additionally, a matchup analysis focused on Estonian inland waters and Baltic Sea coastal waters (n = 22) demonstrated that POLYMER was the most suitable and provided the most amount of unflagged data and C2RCC was not suitable for highly absorbing waters (Alikas et al., 2020).

The contrasting radiometric performance across various regional coastal water bodies (Table 2) demonstrates the challenge of developing a robust atmospheric correction algorithm for global coastal waters. Extensive validation with regional *in situ* data is required to ensure water-leaving retrievals are reliable for input into inversion models to derive accurate water quality components used for research and management. Comparison of performance in different regions can also yield vital information to refine assumptions made in different algorithms, with regards to the inherent optical properties of different regimes.

It is also important to note that there are other atmospheric correction algorithms that were not evaluated in this study that could potentially perform better than the processors applied in this study. For example, Vanhellemont & Ruddick (2021) apply ACOLITE, an atmospheric correction processor developed by the Royal Belgian Institute of Natural Sciences to OLCI imagery of coastal Belgian waters. The adaptation of the 'Dark Spectrum Fitting' approach was shown to provide the best performance in the visible bands in comparison to other atmospheric correction processors including C2RCC and POLYMER. Xue et al. (2019) applied the 6SV atmospheric correction (the vector version of the Second Simulation of the Satellite Signal in the Solar Spectrum correction scheme) to OLCI imagery of Chinese inland waters which demonstrated greater performance than C2RCC and POLYMER.

Schroeder et al. (2022) developed a new coastal atmospheric correction algorithm for OLCI that consists of an ensemble of ANNs and obtained lower mean absolute percentage error and band-averaged bias when compared to C2RCC. This algorithm also provides pixel-based estimation of model inversion uncertainty and sensor noise propagation which is critical to determine the quality with which ocean color products can be estimated from (Schroeder et al., 2022).

Matchup uncertainties

Although radiometric *in situ* measurements are typically used to assess satellite-derived measurements, they can contain variability and uncertainties due to instrument calibration, instrument or ship shadowing, environmental factors, and methods to address surface reflection factors (Zibordi et al., 2012; Shang et al., 2017; Ruddick et al., 2019; Vabson et al., 2019; Tilstone et al., 2020). Quantifying uncertainties from *in situ* measurements is an important step to obtain a reliable fiducial reference measurement and has been performed in previous studies (Vabson et al., 2019; Alikas et al., 2020; Antoine et al., 2021). An intercomparison of *in situ* radiometers and computing an uncertainty budget were not the main objectives for this study and therefore were not included. However, quality control measures were taken to eliminate outliers *in situ* spectra. The agreement between R_{rs} derived from the two *in situ* measurements techniques (Figure 3) provides assurance that measurements were collected with methods that minimize uncertainties. Nonetheless, additional measures to reduce and compute uncertainty would potentially provide higher precision and accuracy of *in situ* R_{rs}.

Radiometric data can also be influenced by viewing and illumination geometries (i.e., solar zenith angle, viewing zenith angle, relative azimuth angle between the Sun and the sensor). This geometry dependence, known as BRDF, aims to characterize how water-leaving radiance varies depending on viewing and illumination angles (Morel et al., 2002). A BRDF correction converts measured water-leaving radiance at any viewing geometry to a conceptual "exact normalized" water-leaving radiance measurement at a nadir direction using values obtained through a look-up table derived by Morel et al. (2002). This correction is based on atmospheric and bio-optical models developed for open ocean waters and requires *a priori* knowledge of chlorophyll *a* concentration. This BRDF correction has been validated against open ocean measurements to show good agreement (Voss et al., 2007); however, it can fail in regions with different atmospheric conditions and confounding inherent optical properties, such as coastal waters. Methods to improve the BRDF correction in coastal waters is an active area of research (Gilerson et al., 2011; Fan et al., 2016); however, due to the current uncertainty, this study did not apply BRDF corrections to *in situ* or satellite-derived R_{rs}. A well performing BRDF correction for coastal waters could potentially improve matchup uncertainties.

Ocean color sensors are pre-calibrated before launch; however, residual calibration errors develop throughout time, requiring a calibration process to ensure satellite retrievals meet water-leaving radiance requirements (i.e., L_w within 5%

uncertainties, Franz et al., 2007; Mobley et al., 2016). This process includes measuring the difference between satellite retrieved TOA radiance to radiance estimated by back propagating high quality *in situ* water leaving radiance to the TOA (Franz et al., 2007; Mobley et al., 2016). Therefore, calibration gain factors are specific to a particular sensor and atmospheric correction algorithm. Gain factors, often termed system vicarious calibration (SVC) gains, are recomputed by conducting a series of matchup comparisons of satellite and *in situ* data and are applied in routine atmospheric correction processing chains. Recommended SVC gains and sources for each atmospheric correction algorithm can be found in Table S2. Publicly available Level 2 OLCI data (BAC) is processed the most current SVC gains (EUMETSAT, 2021). In this study, BAC data was downloaded at different times throughout 2019 to 2021 and contained varying SVC gains which could potentially impact matchup results. It is recommended to download data once to ensure SVC gains are consistent throughout. C2RCC was processed with SVC gains listed in the 2021 EUMETSAT Level 2 report for baseline collection (EUMETSAT, 2021). POLYMER was processed with the default SVC gains hardcoded in POLYMER v4.12 which were updated in 2019[3]. OLCI-A gains were applied by default to OLCI-B, which could potentially impact the accuracy of OLCI-B retrievals since gains should be sensor specific. L2gen was processed with SVC gains derived following procedures of Franz et al. (2007) and Werdell et al. (2007) with the use of a model based on a climatology

[3] POLYMER changelog, 2021-12-17. Accessed online 13 March 2022: https://www.hygeos.com/static/polymer/CHANGELOG.TXT

of chlorophyll at the Marine Optical Buoy (MOBY) site[4]. Operational L2gen SVC gains are derived using open ocean calibration targets and typical maritime atmospheric conditions which may not be representative of a coastal water body (Franz et al., 2007). Therefore, regionally specific vicarious calibration may be necessary in coastal and inland waters and would potentially improve the matchup analysis in this study.

Pixel flagging

Many solutions to atmospheric correction challenges have been operationally implemented in processing schemes; however, there are some issues where satisfactory solutions have not been developed, causing atmospheric correction to fail, or flag pixels as suspect quality such as absorbing aerosols, whitecaps, and the adjacency effect (Frouin et al., 2019). POLYMER produced more clear day spatial coverage than the other atmospheric correction algorithms, likely due to POLYMER being capable of retrieving data under sun-glint conditions. L2gen resulted in the highest amount of flagged pixels, likely due to the restrictive filtering of stray light and sun glint, which has been shown to be the primary quality flags affecting L2gen data quantity (Hu et al., 2019) and the 500 m default land mask. More pixels were flagged in the Upper Bay for all atmospheric correction algorithms, likely due to high reflectance from scattering suspended mineral particles associated with terrestrial

[4] NASA Ocean Color Forum Post. Accessed online 13 March 2022: https://forum.earthdata.nasa.gov/viewtopic.php?f=7&t=1273&sid=5e1209c2e4f00bec7683116cc017d6fb

runoff or errors associated with nonuniform aerosols. Future work should focus on relaxing filtering criteria to observe differences in data quality and quantity.

Management implications

Satellite remote sensing provides water quality observations at spatial and temporal scales that exceed current field monitoring techniques and can be used to enhance the understanding of coastal processes and monitoring of marine resources. Water quality concentrations such as chlorophyll *a* provide indicators of nutrient enrichment and can be used to predict eutrophic conditions (Harding et al., 2019). Anomalous high levels of chlorophyll *a* concentration can also be used to identify harmful algal blooms which are only expected to increase in abundance and severity with climate change (Glibert, 2019; Wolny et al., 2020). Measuring satellite-derived TSM concentrations can provide insight into nutrient loads, salinity, and light levels (Ondrusek et al., 2012).

Satellite remote sensing has been used to study water quality trends in Chesapeake Bay with various ocean color sensors (Harding et al., 2005; Gitelson et al., 2007; Werdell et al., 2009; Wang et al., 2009; Son & Wang, 2012; Ondrusek et al., 2012; Feng et al., 2018; Tomlinson et al., 2019; Wolny et al., 2020; Turner et al., 2021). Several of these studies demonstrate erroneous water-leaving radiance due to the narrowing shape of the estuary, high optical complexity, and increased urban aerosols that confound the atmospheric correction process (Werdell et al., 2009; Son & Wang, 2012). These uncertainties have prompted the use of utilizing R_{rhos} or R_{rc}

instead of R_{rs} to derive water quality parameters (Feng et al., 2018; Tomlinson et al., 2019; Wolny et al., 2020). R_{rhos} is derived by a partial atmospheric correction that corrects TOA radiance for gaseous absorption (mainly ozone) and Rayleigh scattering effects, resulting in Rayleigh corrected reflectance (Feng et al., 2018). However, since this method does not account for aerosol contributions, it is likely that R_{rhos} has aerosol associated uncertainties. Other studies have used custom workflows for MODIS Aqua data based on a threshold value at R_{rs} at 645 nm that switches to SWIR radiances, unavailable with OLCI, in highly turbid waters (Aurin et al., 2013; Turner et al., 2021).

OLCI currently contains the most spectral bands of any ocean color satellite sensor and provides nearly daily coverage at 300 m spatial resolution; characteristics which make it an important sensor for near real time retrieval of ocean color products. A reliable and well validated atmospheric correction algorithm for Chesapeake Bay waters is necessary to utilize the benefits OLCI contains for reliable water quality monitoring. The C2RCC atmospheric correction is the best performing algorithm for Chesapeake Bay waters in this study and agencies can consider incorporating R_{rs} and water quality products produced by C2RCC into operational data portals (i.e., NOAA CoastWatch, NASA OceanColor Web) to help end-users to select the most appropriate and accurate data for their needs.

Future approaches

Atmospheric correction is done on a pixel by pixel basis, which can be computationally intensive due to the large number of pixels contained in multiple satellite images captured by numerous satellite sensors per day. Therefore, algorithms need to be very computationally fast while maintaining accuracy to produce data in near real-time. Algorithms will likely advance and change in accordance with computational power, new ocean color sensors, and innovative ideas, much like they have over the past several decades. There have been several new proposed atmospheric correction algorithms that have the potential to enhance the upcoming Plankton, Aerosol, Cloud, and ocean Ecosystem (PACE) mission, expected to launch in 2024. PACE will host the first hyperspectral ocean color sensor (the Ocean Color Instrument, OCI) which will measure unprecedented TOA hyperspectral radiance from 340 to 890 nm at 5-nm spectral resolution (Werdell et al., 2019). It will also contain seven discrete channels in the SWIR with high radiometric performance to enable atmospheric correction over open ocean and coastal waters (Ibrahim et al., 2019). Ibrahim et al. (2019) describes a multi-band atmospheric correction (MBAC) which utilizes six bands from the NIR and SWIR region as compared to the heritage two-band ratio aerosol correction algorithm (Gordon & Wang, 1994). This multiband approach reduces uncertainty caused by sensor noise using adaptive spectral weights to decrease the contamination of a highly scattering signal (Ibrahim et al., 2019). Preliminary analyses show significant improvement in radiometric retrievals when compared to NASA's two-step atmospheric correction procedure (Ibrahim et al., 2019).

Quantifying polarizing properties of reflected sunlight also has the potential to improve atmospheric correction. Solar radiation is unpolarized when it enters the Earth's atmosphere; as it interacts with aerosols and molecules and refracted at the atmosphere-ocean interface, it can become partially polarized. Studies have shown that polarization of incident light on the sea surface can lead to significant variations in water-leaving radiances (Liu et al., 2017). Accounting for polarization can enhance the contribution of water-leaving reflectance, improve retrievals impacted by sun glint, and better characterize aerosol properties (Frouin et al., 2019). It has also been demonstrated that measuring polarization at multiple viewing angles can significantly improve aerosol radiation properties which can assist in determining appropriate aerosol types or models in atmospheric correction schemes (Frouin et al., 2019). Multi-angle polarimetry is an emerging field of active research and will be included in the upcoming PACE mission to improve atmospheric correction (Werdell et al., 2019).

Conclusions

Coastal and estuarine ecosystems are dynamic environments where tidal forces, sediment resuspension, and riverine and terrigenous inputs continuously work to alter the relative concentrations of optically active constituents. The OLCI sensor can resolve optical data in coastal water bodies at higher spatial (300 m), spectral (21 bands), and temporal (2-day) resolutions than any other operational ocean color satellite sensor. Since water-leaving radiance is at most 10% of the total radiance

OLCI collects, accurate atmospheric correction is essential. The challenge of atmospheric correction is exacerbated in coastal waters due to the optically complex water-leaving reflectance signal and the presence of strongly absorbing aerosols with differing composition and vertical distributions. This study conducted a radiometric evaluation of four different atmospheric correction algorithms compared to a regional *in situ* dataset from Chesapeake Bay waters. A matchup analysis demonstrated that C2RCC had the best performance, particularly in the longer wavelengths. The results from this study both agreed and contrasted with similar studies in regional coastal water bodies, depicting the challenge of developing a robust atmospheric correction algorithm that works in all global coastal waters. Future research will include comparing biophysical products derived from C2RCC processed data to an extensive dataset of *in situ* water quality concentrations collected in the Chesapeake Bay to evaluate performance and analyze water quality variability. This study provides a framework with associated uncertainties and recommendations to utilize OLCI ocean color data to monitor the biogeochemical dynamics in the Chesapeake Bay.

Supplementary Material

Table S1. Details of downloaded OLCI data preprocessed with BAC and processing parameters applied to C2RCC, POLYMER, L2gen atmospheric correction algorithms.

BAC	**C2RCC**	**POLYMER**	**l2gen**
Level 1: Full resolution (EFR), Non-time critical	SNAP c2rcc.olci operator	water_model = PR05 (‘based on Park and Ruddick, 2005’)	brdf_opt = 0 (‘no bidirectional reflectance correction’)

Level 2: Full resolution (WFR), Non-time critical	bbopt, min_abs = 0 (‘don’t include particle absorption’) absorption = ‘bricaud98_aphy’ normalize = 2 (‘Apply wavelength normalization for MERIS and OLCI’)	aer_opt = -3 (‘aerosol mode option is multi-scattering with 2-band model selection and iterative NIR correction’) aer_wave_short = 865 (‘shortest sensor wavelength for aerosol model selection’) aer_wave_long = 1012 (‘longest sensor wavelength for aerosol model selection’)

Table S2. OLCI system vicarious calibration (SVC) gains applied in C2RCC, POLYMER, L2gen processing of Sentinel-3A and -3B OLCI data.

	400	**412.5**	**442.5**	**490**	**510**	**560**	**620**	**665**	**673.8**	**681.3**	**708.8**	**Source**
C2RCC-S3A	0.97546	0.97406	0.97492	0.96890	0.97184	0.97571	0.98001	0.97834	0.97860	0.97908	0.98013	Sentinel-3 OLCI L2 report for baseline collection OL_L2M_003[1]
	753.8	**761.3**	**764.4**	**767.5**	**778.8**	**865**	**885**	**900**	**940**	**1020**		
	0.98522	1.00000	1.00000	1.00000	0.98772	0.98600	0.98657	1.00000	1.00000	0.91316		
	400	**412.5**	**442.5**	**490**	**510**	**560**	**620**	**665**	**673.8**	**681.3**	**708.8**	
C2RCC-S3B	0.99458	0.9901	0.99221	0.9862	0.98898	0.99114	0.99769	0.99684	0.99716	0.99802	0.99782	
	753.8	**761.3**	**764.4**	**767.5**	**778.8**	**865**	**885**	**900**	**940**	**1020**		
	1.00163	1.0000	1.0000	1.0000	1.00259	1.0000	1.00089	1.0000	1.0000	0.94064		
	400	**412**	**443**	**490**	**510**	**560**	**620**	**665**	**674**	**681**	**709**	As provided by POLYMER v4.12
Polymer-S3A & S3B	1.0470	1.0470	1.0400	1.0560	1.0320	1.0330	1.0180	1.0310	1.0279	1.0291	1.0255	
	754	**760**	**764**	**767**	**779**	**865**	**885**	**900**	**940**	**1020**		
	1.0317	1.0271	1.0245	1.0224	1.0224	1.0262	1.0235	1.0235	1.0235	1.0235		
	400	**412**	**442**	**490**	**510**	**560**	**620**	**665**	**674**	**681**	**709**	Pers. Comm., Sean Bailey, NASA Forum post[2]
L2gen-S3A	0.959728	0.972313	0.97156	0.96916	0.97636	0.97951	0.97705	0.975398	0.97344	0.97597	1.00562	
	754	**761**	**764**	**768**	**779**	**865**	**885**	**900**	**940**	**1020**		
	0.9829	1.0000	1.0000	1.0000	0.98989	1.0000	1.01815	1.0000	1.0000	1.0000		
	400	**412**	**442**	**490**	**510**	**560**	**620**	**665**	**674**	**681**	**709**	
L2gen-S3B	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	
	754	**761**	**764**	**768**	**779**	**865**	**885**	**900**	**940**	**1020**		
	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000		

[1] Available online at https://www-cdn.eumetsat.int/files/2021-04/Sentinel-3%20OLCI%20L2%20report%20for%20baseline%20collection%20OL_L2M_003_2B.pdf, accessed 13 March 2022

[2] Available online at https://forum.earthdata.nasa.gov/viewtopic.php?f=7&t=1273&sid=41a6bf1f32d110c330ffe7a1da00c10b#confirm_external_link-modal, accessed 13 March 2022

Table S3. Recommended flags applied to each atmospheric correction algorithm.

	Recommended Flags	Source
BAC	Quality flags (WQSF_lsb): CLOUD, CLOUD_AMBIGUOUS, CLOUD_MARGIN, INVALID, COSMETIC, SATURATED, SUSPECT, HISOLZEN, HIGHGLINT, SNOW_ICE, AC_FAIL, WHITECAPS + Science flags (WQSF_msb): RWNEG_O2, RWNEG_O3, RWNEG_O4, RWNEG_O5, RWNEG_O6, RWNEG_O7, RWNEG_O8 + IDEPIX Flags (pixel_classif_flags): IDEPIX_INVALID, IDEPIX_CLOUD, IDEPIX_CLOUD_AMBIGUOUS, IDEPIX_CLOUD_SURE, IDEPIX_CLOUD_BUFFER, IDEPIX_CLOUD_SHADOW, IDEPIX_SNOW_ICE, IDEPIX_BRIGHT, IDEPIX_WHITE, IDEPIX_COASTLINE, IDEPIX_LAND	EUMETSAT Sentinel-3 Product Notice – OLCI Level-2 Ocean Colour (2021)[1], Kratzer & Plowey, 2021
C2RCC	Quality flags (WQSF_lsb): CLOUD, CLOUD_AMBIGUOUS, CLOUD_MARGIN, INVALID, COSMETIC, SATURATED, SUSPECT, HISOLZEN, HIGHGLINT, SNOW_ICE + IDEPIX flags	EUMETSAT Sentinel-3 OLCI Matchup Protocols (2021)[2]
POLYMER	Quality flags (bitmask): LAND, CLOUD_BASE, L1_INVALID, NEGATIVE_BB, OUT_OF_BOUNDS, EXCEPTION, THICK_AEROSOL, HIGH_AIR_MASS, EXTERNAL_MASK (bitmask & 1023 = 0) + IDEPIX flags	POLYMER v4.13 README.md
l2gen	Quality flags (l2_flags): ATMFAIL, LAND, HIGLINT, HILT, HISATZEN, STRAYLIGHT,	NASA Ocean Color Web[3]

CLDICE, COCCOLITH, HISOLZEN,
LOWLW, CHLFAIL, NAVWARN,
MAXAERITER, CHLWARN, ATMWARN,
NAVFAIL
+
IDEPIX flags

---

[1] Available online at https://www.eumetsat.int/media/48139, accessed 13 March 2022
[2] Available online at https://www.eumetsat.int/media/44087, accessed 13 March 2022
[3] Available online at https://oceancolor.gsfc.nasa.gov/atbd/ocl2flags/, accessed 13 March 2022

References

Ahmad, Z., Franz, B. A., McClain, C. R., Kwiatkowska, E. J., Werdell, J., Shettle, E. P., & Holben, B. N. 2010. New aerosol models for the retrieval of aerosol optical thickness and normalized water-leaving radiances from the SeaWiFS and MODIS sensors over coastal regions and open oceans. *Applied optics*, *49*(29), 5545-5560.

Ahn, Y. H., Ryu, J. H., and Moon, J. E. 1999. Development of Redtide & Water Turbidity Algorithms Using Ocean Color Satellite. KORDI Report No. BSPE 98721-00-1224-01. Seoul, Korea: KORDI.

Alikas, K., Ansko, I., Vabson, V., Ansper, A., Kangro, K., Uudeberg, K., & Ligi, M. 2020. Consistency of Radiometric Satellite Data over Lakes and Coastal Waters with Local Field Measurements. *Remote Sensing*, *12*(4), 616.

Antoine, D., & Morel, A. 1999. A multiple scattering algorithm for atmospheric correction of remotely sensed ocean colour (MERIS instrument): principle and implementation for atmospheres carrying various aerosols including absorbing ones. *International Journal of Remote Sensing*, *20*(9), 1875-1916.

Antoine, D., Slivkoff, M., Klonowski, W., Kovach, C., & Ondrusek, M. 2021. Uncertainty assessment of unattended above-water radiometric data collection from research vessels with the Dynamic Above-water Radiance (L) and Irradiance (E) Collector (DALEC). *Optics Express*, *29*(3), 4607-4631.

Aurin, D., Mannino, A. and Franz, B., 2013. Spatially resolving ocean color and sediment dispersion in river plumes, coastal systems, and continental shelf waters. *Remote Sensing of Environment*, *137*, pp.212-225.

Bailey, S. W., Franz, B. A., & Werdell, P. J. 2010. Updated NIR water-leaving radiance estimation for ocean color data processing. *Optics Express*, *18*(7), 7521-7527.

Barker, K., Mazeran, C., Lerebourg, C., Bouvet, M., Antoine, D., Ondrusek, M., Zibordi, G. and Lavender, S. 2008. MERMAID: The MEris MAtchup *in-situ* database. In *Proceedings of the 2nd (A) ATSR and MERIS Workshop, Frascati, Italy*.

Brockmann, C., Doerffer, R., Peters, M., Kerstin, S., Embacher, S., & Ruescas, A.

2016. Evolution of the C2RCC neural network for Sentinel 2 and 3 for the retrieval of ocean colour products in normal and extreme optically complex waters. In *Living Planet Symposium* (Vol. 740, p. 54).
Doerffer, R., & Schiller, H. 2007. The MERIS Case 2 water algorithm. *International Journal of Remote Sensing*, *28*(3-4), 517-535.
EUMETSAT. 2018. Sentinel-3 OLCI Marine User Handbook. Available online at: https://www.eumetsat.int/website/home/Data/TechnicalDocuments/index.html?section=copernicus#copernicus, accessed 17 January 2020.
EUMETSAT. 2021. Sentinel-3 OLCI L2 report for baseline collection OL_L2M_003. Available online at: https://www.eumetsat.int/media/47794, accessed 26 October 2021.
Fan, Y., Li, W., Voss, K. J., Gatebe, C. K., & Stamnes, K. 2016. Neural network method to correct bidirectional effects in water-leaving radiance. *Applied optics*, *55*(1), 10-21.
Feng, L., Hou, X., Li, J., & Zheng, Y. 2018. Exploring the potential of Rayleigh-corrected reflectance in coastal and inland water applications: A simple aerosol correction method and its merits. *ISPRS Journal of Photogrammetry and Remote Sensing*, *146*, 52-64.
Franz, B. A., Bailey, S. W., Werdell, P. J., & McClain, C. R. 2007. Sensor-independent approach to the vicarious calibration of satellite ocean color radiometry. *Applied Optics*, *46*(22), 5068-5082.
Froidefond, J. M., & Ouillon, S. 2005. Introducing a mini-catamaran to perform reflectance measurements above and below the water surface. *Optics Express*, *13*(3), 926-936.
Frouin, R.J., Franz, B.A., Ibrahim, A., Knobelspiesse, K., Ahmad, Z., Cairns, B., Chowdhary, J., Dierssen, H.M., Tan, J., Dubovik, O. and Huang, X. 2019. Atmospheric correction of satellite ocean-color imagery during the PACE era. *Frontiers in earth science*, *7*, 145.
Giannini, F., Hunt, B. P., Jacoby, D., & Costa, M. 2021. Performance of OLCI Sentinel-3A satellite in the Northeast Pacific coastal waters. *Remote Sensing of Environment*, *256*, 112317.
Gilerson, A., Hlaing, S., Harmel, T., Tonizzo, A., Arnone, R., Weidemann, A., & Ahmed, S. 2011. Bidirectional reflectance function in coastal waters: modeling and validation. In *Remote Sensing of the Ocean, Sea Ice, Coastal Waters, and Large Water Regions 2011* (Vol. 8175, p. 81750O). International Society for Optics and Photonics.
Gitelson, A. A., Schalles, J. F., & Hladik, C. M. 2007. Remote chlorophyll-a retrieval in turbid, productive estuaries: Chesapeake Bay case study. *Remote Sensing of Environment*, *109*(4), 464-472.
Gitelson, A.A., Dall'Olmo, G., Moses, W., Rundquist, D.C., Barrow, T., Fisher, T.R., Gurlin, D., Holz, J. 2008. A simple semi-analytical model for remote estimation of chlorophyll-*a* in turbid waters: Validation. *Remote Sensing of Environment*, *112*(9), 3582-3593.

Glibert, P. M. 2020. Harmful algae at the complex nexus of eutrophication and climate change. *Harmful Algae*, *91*, 101583.

Gons, H. J., Rijkeboer, M., & Ruddick, K. G. 2002. A chlorophyll-retrieval algorithm for satellite imagery (Medium Resolution Imaging Spectrometer) of inland and coastal waters. *Journal of Plankton Research*, *24*(9), 947-951.

Gordon, H. R. 1978. Removal of atmospheric effects from satellite imagery of the oceans. *Applied Optics*, *17*(10), 1631-1636.

Gordon, H. R., & Wang, M. 1994. Retrieval of water-leaving radiance and aerosol optical thickness over the oceans with SeaWiFS: a preliminary algorithm. *Applied optics*, *33*(3), 443-452.

Gossn, J. I., Ruddick, K. G., & Dogliotti, A. I. 2019. Atmospheric correction of OLCI imagery over extremely turbid waters based on the red, NIR and 1016 nm bands and a new baseline residual technique. *Remote Sensing*, *11*(3), 220.

Goyens, C., Jamet, C., & Schroeder, T. 2013. Evaluation of four atmospheric correction algorithms for MODIS-Aqua images over contrasted coastal waters. *Remote Sensing of Environment*, *131*, 63-75.

Harding Jr, L. W., Magnuson, A., & Mallonee, M. E. 2005. SeaWiFS retrievals of chlorophyll in Chesapeake Bay and the mid-Atlantic bight. *Estuarine, Coastal and Shelf Science*, *62*(1-2), 75-94.

Hieronymi, M., Müller, D., & Doerffer, R. 2017. The OLCI Neural Network Swarm (ONNS): a bio-geo-optical algorithm for open ocean and coastal waters. *Frontiers in Marine Science*, *4*, 140.

Hu, C., Barnes, B. B., Feng, L., Wang, M., & Jiang, L. 2019. On the interplay between ocean color data quality and data quantity: Impacts of quality control flags. *IEEE Geoscience and Remote Sensing Letters*, *17*(5), 745-749.

Ibrahim, A., Franz, B. A., Ahmad, Z., & Bailey, S. W. 2019. Multiband atmospheric correction algorithm for ocean color retrievals. *Frontiers in Earth Science*, *7*, 116.

IOCCG. 2000. Remote Sensing of Ocean Colour in Coastal, and Other Optically-Complex, Waters. Sathyendranath, S. (ed.), Reports of the International Ocean-Colour Coordinating Group, No. 3, IOCCG, Dartmouth, Canada.

IOCCG. 2006. Remote Sensing of Inherent Optical Properties: Fundamentals, Tests of Algorithms, and Applications. Lee, Z.-P. (ed.), Reports of the International Ocean-Colour Coordinating Group, No. 5, IOCCG, Dartmouth, Canada.

IOCCG. 2010. Atmospheric Correction for Remotely-Sensed Ocean-Colour Products. Wang, M. (ed.), Reports of the International Ocean-Colour Coordinating Group, No. 10, IOCCG, Dartmouth, Canada.

Jolliff, J. K., Kindle, J. C., Shulman, I., Penta, B., Friedrichs, M. A., Helber, R., & Arnone, R. A. 2009. Summary diagrams for coupled hydrodynamic-ecosystem model skill assessment. *Journal of Marine Systems*, *76*(1-2), 64-82.

Kemp, W.M., Boynton, W.R., Adolf, J.E., Boesch, D.F., Boicourt, W.C., Brush, G., Cornwell, J.C., Fisher, T.R., Glibert, P.M., Hagy, J.D. and Harding, L.W. 2005. Eutrophication of Chesapeake Bay: historical trends and ecological interactions. *Marine ecology progress series*, *303*, 1-29.

Kratzer, S., & Plowey, M. 2021. Integrating mooring and ship-based data for improved validation of OLCI chlorophyll-*a* products in the Baltic Sea. *International Journal of Applied Earth Observation and Geoinformation*, *94*, 102212.

Kuchinke, C. P., Gordon, H. R., Harding Jr, L. W., & Voss, K. J. 2009. Spectral optimization for constituent retrieval in Case 2 waters II: Validation study in the Chesapeake Bay. *Remote Sensing of Environment*, *113*(3), 610-621.

Kyryliuk, D., & Kratzer, S. 2019. Evaluation of Sentinel-3A OLCI products derived using the Case-2 Regional CoastColour processor over the Baltic Sea. *Sensors*, *19*(16), 3609.

Lavender, S. J., Pinkerton, M. H., Moore, G. F., Aiken, J., & Blondeau-Patissier, D. 2005. Modification to the atmospheric correction of SeaWiFS ocean colour images over turbid waters. *Continental Shelf Research*, *25*(4), 539-555.

Lee, Z. P., Du, K. P., & Arnone, R. 2005. A model for the diffuse attenuation coefficient of downwelling irradiance. *Journal of Geophysical Research: Oceans*, *110*(C2).

Lee, Z., Pahlevan, N., Ahn, Y. H., Greb, S., & O'Donnell, D. 2013. Robust approach to directly measuring water-leaving radiance in the field. *Applied Optics*, *52*(8), 1693-1701.

Lee, Z., Wei, J., Shang, Z., Garcia, R., Dierssen, H. M., Ishizaka, J., & Castagna, A. 2019. On-water radiometry measurements: skylight-blocked approach and data processing. *Appendix to Protocols for Satellite Ocean Colour Data Validation: In Situ Optical Radiometry. IOCCG Ocean Optics and Biogeochemistry Protocols for Satellite Ocean Colour Sensor Validation, Volume 3.0*, 7.

Liu, J., He, X., Liu, J., Bai, Y., Wang, D., Chen, T., Wang, Y. and Zhu, F., 2017. Polarization-based enhancement of ocean color signal for estimating suspended particulate matter: radiative transfer simulations and laboratory measurements. Optics express, 25(8), pp.A323-A337.

Mannino, A., Russ, M. E., & Hooker, S. B. 2008. Algorithm development and validation for satellite-derived distributions of DOC and CDOM in the US Middle Atlantic Bight. *Journal of Geophysical Research: Oceans*, *113*(C7).

Miller, W. D., Kimmel, D., & Harding, L. W., Jr. 2006. Predicting spring discharge of the Susquehanna River from a winter synoptic climatology for the eastern United States. *Water Resources Research*, 42.

Mobley, C. D., & Mobley, C. D. 1994. *Light and water: radiative transfer in natural waters*. Academic press.

Mobley, C. D., Werdell, J., Franz, B., Ahmad, Z., & Bailey, S. 2016. Atmospheric correction for satellite ocean color radiometry.

Moore, G. F., Aiken, J., & Lavender, S. J. 1999. The atmospheric correction of water colour and the quantitative retrieval of suspended particulate matter in Case II waters: application to MERIS. *International Journal of Remote Sensing*, *20*(9), 1713-1733.

Moore, G., C. Mazeran and J.-P. Huot. 2017. Case II.S Bright Pixel Atmospheric Correction. Algorithm Theoretical. Basis Document(ATBD) 2.6; MERIS ATBD-ESA Earth Online: Frascati, Italy, 2017; p. 82.

Mograne, M. A., Jamet, C., Loisel, H., Vantrepotte, V., Mériaux, X., & Cauvin, A. 2019. Evaluation of five atmospheric correction algorithms over French optically-complex waters for the Sentinel-3A OLCI Ocean Color Sensor. *Remote Sensing*, *11*(6), 668.

Morel, A., Antoine, D., & Gentili, B. 2002. Bidirectional reflectance of oceanic waters: accounting for Raman emission and varying particle scattering phase function. *Applied Optics*, *41*(30), 6289-6306.

Moses, W. J., Saprygin, V., Gerasyuk, V., Povazhnyy, V., Berdnikov, S., & Gitelson, A. A. 2019. OLCI-based NIR-red models for estimating chlorophyll-a concentration in productive coastal waters—a preliminary evaluation. *Environmental Research Communications*, *1*(1), 011002.

Mouw, C.B., Greb, S., Aurin, D., DiGiacomo, P.M., Lee, Z., Twardowski, M., Binding, C., Hu, C., Ma, R., Moore, T. and Moses, W., 2015. Aquatic color radiometry remote sensing of coastal and inland waters: Challenges and recommendations for future satellite missions. *Remote Sensing of Environment,* 160, 15-30.

Nechad, B., Ruddick, K. G., & Park, Y. 2010. Calibration and validation of a generic multisensor algorithm for mapping of total suspended matter in turbid waters. *Remote Sensing of Environment*, *114*(4), 854-866.

Ondrusek, M., Stengel, E., Kinkade, C. S., Vogel, R. L., Keegstra, P., Hunter, C., & Kim, C. 2012. The development of a new optical total suspended matter algorithm for the Chesapeake Bay. *Remote Sensing of Environment*, *119*, 243-254.

Ondrusek, M. E., Stengel, E., Rella, M. A., Goode, W., Ladner, S., & Feinholz, M. 2014. Validation of ocean color sensors using a profiling hyperspectral radiometer. In *Ocean Sensing and Monitoring VI* (Vol. 9111, p. 91110Y). International Society for Optics and Photonics.

O'Reilly, J. E., Maritorena, S., Mitchell, B. G., Siegel, D. A., Carder, K. L., Garver, S. A., Kahru, M., & McClain, C. 1998. Ocean color chlorophyll algorithms for SeaWiFS. *Journal of Geophysical Research: Oceans*, *103*(C11), 24937-24953.

Ruddick, K.G., De Cauwer, V., Park, Y.J. and Moore, G., 2006. Seaborne measurements of near infrared water-leaving reflectance: The similarity spectrum for turbid waters. *Limnology and Oceanography*, *51*(2), pp.1167-1179.

Pereira-Sandoval, M., Ruescas, A., Urrego, P., Ruiz-Verdú, A., Delegido, J., Tenjo,

C., Soria-Perpinyà, X., Vicente, E., Soria, J. and Moreno, J., 2019. Evaluation of atmospheric correction algorithms over Spanish inland waters for sentinel-2 multi spectral imagery data. Remote Sensing, 11(12), p.1469.
Renosh, P. R., Doxaran, D., Keukelaere, L. D., & Gossn, J. I. 2020. Evaluation of atmospheric correction algorithms for sentinel-2-msi and sentinel-3-olci in highly turbid estuarine waters. *Remote Sensing*, *12*(8), 1285.
Ruddick, K. G., Ovidio, F., & Rijkeboer, M. 2000. Atmospheric correction of SeaWiFS imagery for turbid coastal and inland waters. *Applied optics*, *39*(6), 897-912.
Ruddick, K.G., Voss, K., Boss, E., Castagna, A., Frouin, R., Gilerson, A., Hieronymi, M., Johnson, B.C., Kuusk, J., Lee, Z. and Ondrusek, M., 2019. A review of protocols for fiducial reference measurements of water-leaving radiance for validation of satellite remote-sensing data over water. Remote Sensing, 11(19), p.2198.
Sathyendranath, S., Brewin, R.J., Brockmann, C., Brotas, V., Calton, B., Chuprin, A., Cipollini, P., Couto, A.B., Dingle, J., Doerffer, R. and Donlon, C., 2019. An ocean-colour time series for use in climate studies: the experience of the ocean-colour climate change initiative (OC-CCI). *Sensors*, *19*(19), p.4285.
Schroeder, T., Behnert, I., Schaale, M., Fischer, J., & Doerffer, R. 2007. Atmospheric correction algorithm for MERIS above case-2 waters. *International Journal of Remote Sensing*, *28*(7), 1469-1486.
Schroeder, T., Schaale, M., Lovell, J. and Blondeau-Patissier, D., 2022. An ensemble neural network atmospheric correction for Sentinel-3 OLCI over coastal waters providing inherent model uncertainty estimation and sensor noise propagation. *Remote Sensing of Environment*, 270, p.112848.
Shang, Z., Lee, Z., Dong, Q., & Wei, J. 2017. Self-shading associated with a skylight-blocked approach system for the measurement of water-leaving radiance and its correction. *Applied optics*, *56*(25), 7033-7040.
Shettle, E. P., & Fenn, R. W. 1976. Models of the atmospheric aerosols and their optical properties. *In AGARD Opt. Propagation in the Atmosphere 16 p (SEE N76-29815 20-46*.
Shi, W., & Wang, M. 2009. An assessment of the black ocean pixel assumption for MODIS SWIR bands. *Remote Sensing of Environment*, *113*(8), 1587-1597.
Siegel, D. A., Wang, M., Maritorena, S., & Robinson, W. 2000. Atmospheric correction of satellite ocean color imagery: the black pixel assumption. *Applied optics*, *39*(21), 3582-3591.
Son, S., & Wang, M. 2012. Water properties in Chesapeake Bay from MODIS-Aqua measurements. *Remote Sensing of Environment*, *123*, 163-174.
Soppa, M. A., Silva, B., Steinmetz, F., Keith, D., Scheffler, D., Bohn, N., & Bracher, A. 2021. Assessment of Polymer Atmospheric Correction Algorithm for Hyperspectral Remote Sensing Imagery over Coastal Waters. *Sensors*, *21*(12), 4125.
Spyrakos, E., O'donnell, R., Hunter, P.D., Miller, C., Scott, M., Simis, S.G., Neil, C.,

Barbosa, C.C., Binding, C.E., Bradt, S. and Bresciani, M., 2018. Optical types of inland and coastal waters. *Limnology and Oceanography*, 63(2), pp.846-870.
Steinmetz, F., Deschamps, P. Y., & Ramon, D. 2011. Atmospheric correction in presence of sun glint: application to MERIS. *Optics express*, *19*(10), 9783-9800.
Steinmetz, F., & Ramon, D. 2018. Sentinel-2 MSI and Sentinel-3 OLCI consistent ocean colour products using POLYMER. In *Remote Sensing of the Open and Coastal Ocean and Inland Waters* (Vol. 10778, p. 107780E). International Society for Optics and Photonics.
Stumpf, R. P., Arnone, R. A., Gould, R. W., Jr., Martinolich, P. M., & Ransibrahmanakul, V. 2003. A partially coupled ocean-atmosphere model for retrieval of water-leaving radiance from seawifs in coastal waters. NASA Tech. Memo. 206892, National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, MD
Taylor, K. E. 2001. Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres*, *106*(D7), 7183-7192.
Tilstone, G., Dall'Olmo, G., Hieronymi, M., Ruddick, K., Beck, M., Ligi, M., Costa, M., D'alimonte, D., Vellucci, V., Vansteenwegen, D. and Bracher, A. 2020. Field intercomparison of radiometer measurements for ocean colour validation. *Remote Sensing*, *12*(10), 1587.
Tomlinson, M. C., Stumpf, R. P., & Vogel, R. L. 2019. Approximation of diffuse attenuation, Kd, for MODIS high-resolution bands. *Remote Sensing Letters*, *10*(2), 178-185.
Turner, J.S., Friedrichs, C.T. and Friedrichs, M.A., 2021. Long-term Trends in Chesapeake Bay Remote Sensing Reflectance: Implications for Water Clarity. *Journal of Geophysical Research: Oceans*, e2021JC017959.
Tzortziou, M., Subramaniam, A., Herman, J. R., Gallegos, C. L., Neale, P. J., & Harding Jr, L. W. 2007. Remote sensing reflectance and inherent optical properties in the mid Chesapeake Bay. *Estuarine, Coastal and Shelf Science*, *72*(1-2), 16-32.
Vabson, V., Kuusk, J., Ansko, I., Vendt, R., Alikas, K., Ruddick, K., Ansper, A., Bresciani, M., Burmester, H., Costa, M. and D'Alimonte, D., 2019. Field intercomparison of radiometers used for satellite validation in the 400–900 nm range. Remote Sensing, 11(9), p.1129.
Vanhellemont, Q., & Ruddick, K. 2021. Atmospheric correction of Sentinel-3/OLCI data for mapping of suspended particulate matter and chlorophyll-a concentration in Belgian turbid coastal waters. *Remote Sensing of Environment*, *256*, 112284.
Wang, M., & Shi, W. 2007. The NIR-SWIR combined atmospheric correction approach for MODIS ocean color data processing. *Optics Express*, *15*(24), 15722-15733.

Wang, M. 2007. Remote sensing of the ocean contributions from ultraviolet to near-infrared using the shortwave infrared bands: simulations. *Applied optics*, *46*(9), 1535-1547.

Werdell, P. J., Franz, B. A., & Bailey, S. W. 2010. Evaluation of shortwave infrared atmospheric correction for ocean color remote sensing of Chesapeake Bay. *Remote Sensing of Environment*, *114*(10), 2238-2247.

Werdell, P.J., Behrenfeld, M.J., Bontempi, P.S., Boss, E., Cairns, B., Davis, G.T., Franz, B.A., Gliese, U.B., Gorman, E.T., Hasekamp, O. and Knobelspiesse, K.D., 2019. The Plankton, Aerosol, Cloud, ocean Ecosystem mission: status, science, advances. *Bulletin of the American Meteorological Society*, *100*(9), pp.1775-1794.

Wolny, J. L., Tomlinson, M. C., Schollaert Uz, S., Egerton, T. A., McKay, J. R., Meredith, A., Reece, K.S., Scott, G.P., & Stumpf, R. P. 2020. Current and future remote sensing of harmful algal blooms in the Chesapeake Bay to support the shellfish industry. *Frontiers in Marine Science*, *7*, 337.

Xue, K., Ma, R., Shen, M., Li, Y., Duan, H., Cao, Z., Wang, D. and Xiong, J., 2020. Variations of suspended particulate concentration and composition in Chinese lakes observed from Sentinel-3A OLCI images. Science of the Total Environment, 721, p.137774.

Zheng, G., & DiGiacomo, P. M. 2017. Uncertainties and applications of satellite-derived coastal water quality products. *Progress in oceanography*, *159*, 45-72.

Zibordi, G., Holben, B. N., Talone, M., D'Alimonte, D., Slutsker, I., Giles, D. M., & Sorokin, M. G. 2020. Advances in the ocean color component of the Aerosol Robotic Network (AERONET-OC). *Journal of Atmospheric and Oceanic Technology*.

Chapter 3: Interactions between water quality and optical water types in Chesapeake Bay

Introduction

The Chesapeake Bay, located in the Mid-Atlantic region, is the largest estuary in the United States. The Bay is a coastal plain estuary, running approximately 300 km in length from north to south with a mean depth of 6.5 m. The Chesapeake Bay has a 164,200 km^2 watershed and an 18,800 km shoreline, resulting in a 14:1 land to water ratio and a close connection with the surrounding land (Hagy et al., 2002; Kemp et al., 2005; Harding et al., 2016). Historic human population growth and associated land-use changes in the watershed (predominantly forested to agriculture and urban) has increased nutrient loading to the Bay (Kemp et al., 2005). This nutrient enrichment has increased phytoplankton biomass including harmful algal species, reduced water clarity, enhanced dissolved oxygen depletion in bottom waters, and gradually shifted the foodweb from benthic to predominantly planktonic (Cooper & Brush, 1993; Kemp et al., 2005).

Given the level of concern about estuarine eutrophication from scientists, managers, and policy makers, a cooperative effort was formed in the 1980s to monitor and manage Chesapeake Bay water quality (Boesch et al., 2001). This effort eventually led to the implementation of the Chesapeake Bay Agreement and consequent federally mandated nutrient load reductions (e.g., Total Maximum Daily Loads, TMDLs) (Chesapeake Bay Program, 2014). Targeting and assessing the progress of nutrient reduction efforts requires knowledge of water quality dynamics and trends, which led to the development of the Chesapeake Bay Program (CBP) Water Quality Monitoring Program. Since 1984, CBP has been overseeing regular (bi-monthly or monthly) monitoring of physical, chemical, and biological parameters throughout the year at more than 130 fixed stations throughout the Bay's mainstem and tributaries. This extensive Bay-wide long-term dataset has contributed to a wealth of knowledge about the temporal and spatial trends in water quality and is used to parameterize numerical modeling to help inform management decisions (Xu & Hood, 2006; Harding et al., 2016; Testa et al., 2019; Harding et al., 2019; Murphy et al., 2022). However, *in situ* data are limited by spatial coverage, cruise frequency, and methods changes, and can rapidly change by dynamic coastal processes, which can hinder the ability to gain a comprehensive understanding of the spatial and temporal patterns in water quality dynamics (Cloern & Jassby 2010).

Ocean color satellite remote sensing provides high frequency measurements with fine spatial resolution that can be used to quantify physical and biogeochemical processes in aquatic ecosystems at large scales (IOCCG, 2006). A satellite measures light emerging from the surface of water which is altered by the type and concentration of dissolved and particulate constituents absorbing and scattering light in water, namely phytoplankton pigments, chromophoric dissolved organic matter (CDOM), inorganic particles, and water itself. Remote sensing reflectance (R_{rs}) in the visible part of the spectrum is proportional to the backscattering coefficient (b_b) and inversely proportional to the backscattering plus absorption coefficient ($b_b + a$ where

$a >> b_b$); these coefficients are referred to as inherent optical properties (IOPs). Therefore, the resultant shape and magnitude of R_{rs} (e.g., color) is mainly controlled by IOPs (e.g., b_b/a spectral ratio). In open ocean waters, absorption and backscattering is primarily dominated by low concentrations of phytoplankton and its co-varying particulate and dissolved materials where R_{rs} peaks in the blue portion of the visible spectrum due to high scattering of water molecules (Roesler & Perry, 1995). In estuaries, light absorption by CDOM and phytoplankton pigments combined with absorption and scattering by suspended particles shift the peak R_{rs} values in the green and red portions of the visible spectrum (Roesler & Perry, 1995). When collected in sediment-rich waters, R_{rs} tends to increase in magnitude and peak in the yellow to red wavelengths due to the highly scattering properties of inorganic particles (Gitelson et al., 2007).

R_{rs} spectral variability can be used to infer underlying water quality constituents using algorithms that exploit bio-optical theory to relate R_{rs} to IOPs (IOCCG, 2006; Werdell et al., 2018). Empirical algorithms relate changes in R_{rs} at various wavelengths (e.g., band-ratios) to field observations of a specific water quality parameter (O'Reiley et al., 1999). These algorithms work reasonably well in open ocean waters where changes in spectra are mainly due from phytoplankton; however, the highly variable proportion of IOPs in coastal waters can jeopardize performance (Harding et al., 2005; Zheng & DiGiacomo, 2017). Many empirical algorithms developed for coastal waters utilize red and near infrared (NIR) wavelengths since they are known to be less sensitive to CDOM and non-algal particles (NAP) (Gitelson et al., 2010). Additionally, regional or class-specific algorithms have been developed to improve water quality retrievals in optically complex waters. Moore et al. (2014) utilized a classification scheme that could select and blend type-specific bio-optical algorithms between optically complex waters. A classification procedure that partitions optical water types (OWTs) can be related to the underlying physical and biological processes to provide information on the spatio-temporal dynamics of optically active constituents in a water body (Spyrakos et al., 2018). An OWT classification provides more comprehensive information than single variables and can be a powerful diagnostic tool for extensive aquatic ecosystem monitoring.

This study analyzes long-term *in situ* water quality data in conjunction with remotely sensed data collected in Chesapeake Bay. The purpose of this work is to 1) synthesize long-term spatial and temporal trends *in situ* water quality data (2000 - 2022), 2) apply a spectral clustering classification of satellite R_{rs} data (2016 - 2022) to identify dominant OWTs in Chesapeake Bay, 3) investigate how water quality is constrained within and across OWTs, and 4) evaluate empirical bio-optical algorithms within OWTs. The flow of this paper is as follows. First, *in situ* water quality data are comprehensively analyzed through space and time to discuss seasonal and interannual variability. Next, a machine learning statistical clustering technique is applied to R_{rs} data to identify and map dominant OWTs (i.e., color) in the Chesapeake Bay. Subsequently, *in situ* water quality data are paired in space and time with OWTs to assess the degree to which water quality parameters are constrained within a given

OWT. This data is used to evaluate seven empirical bio-optical algorithms designed to estimate concentrations of chlorophyll *a* from measurements of R_{rs} in optically complex waters to select a dynamic switching algorithm for an improved blended chlorophyll product. We end the paper with management considerations, describing how the results of this study can aid in the interpretation of water quality trends, assist in water quality algorithm selection, and potentially be assimilated into biogeochemical models.

Methods

In situ data sources

In situ Chesapeake Bay water quality data collected at mainstem and tributary stations (Figure 1) from 2000 to 2022 was obtained from the Chesapeake Bay Program (CBP) Tidal Water Quality Database[5]. The following parameters were downloaded: salinity, total nitrogen (TN), total phosphorus (TP), particulate nitrogen (PN), particulate phosphorus (PP), active chlorophyll *a* (CHLA), total suspended solids (TSS), fixed suspended solids (FSS), volatile suspended solids (VSS), light attenuation (KD), and Secchi depth (SECCHI). More information on measurement methods used can be found in the CBP Database User Guide (Chesapeake Bay Program, 2012). If either VSS or FSS were absent, we calculated these values by

[5] Chesapeake Bay Program Water Quality Database (1984-present), accessed online January 21, 2022: https://www.chesapeakebay.net/what/downloads/cbp-water-quality-database-1984-present

subtracting one from TSS. Chl:C was calculated by dividing VSS multiplied by 1/2 by CHLA, assuming that half of VSS is due to phytoplankton carbon. The stoichiometric N:P Redfield ratio was estimated by dividing molar concentrations of PN by the molar concentration of PP, though these data should be interpreted with caution in a strictly Redfield sense due to the presence of particulate N and P associated with inorganic sediments. Z_{SD} is analyzed in terms of the inverse of Z_{SD} to make values more comparable to K_d. Data were filtered to obtain surface measurements (< 1 m) and remove any outlier or problem values (e.g., ignoring values with 'Qualifier' or 'Problem' value).

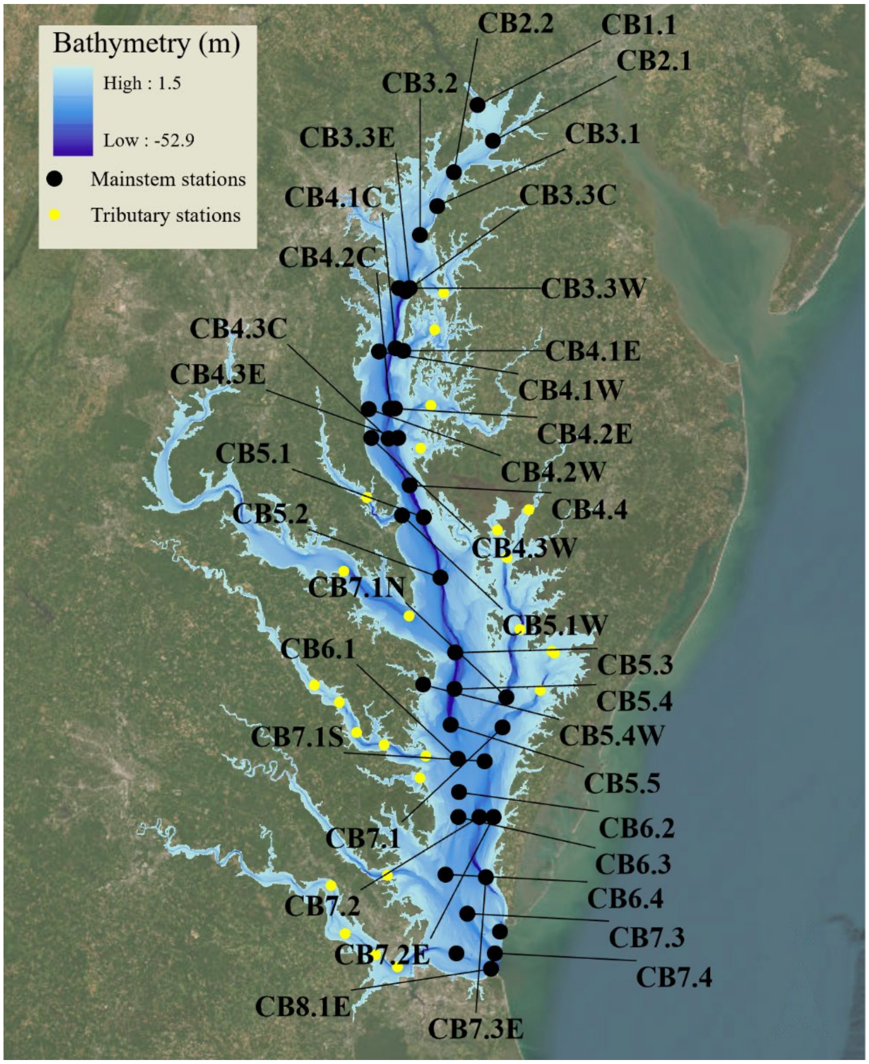


Figure 1. Bathymetric map of Chesapeake Bay (NOAA, 1994) with Chesapeake Bay Program mainstem (black points) and tributary (yellow points) tidal water quality monitoring stations analyzed in this study.

In situ water quality data synthesis

In order to visualize spatiotemporal trends in water quality, monthly and annual climatologies and standardized anomalies were calculated for each water quality parameter (C) collected at 41 mainstem stations (e.g., stations that start with

‘CB’). First, at each station (*s*), water quality data was averaged within a given month (*m*). Station specific monthly climatologies ($MC_{s,m}$) were derived by calculating the mean of monthly data for the period 2000 to 2022. Station specific monthly standardized anomalies ($SA_{s,m}$) were calculated as:

$$SA_{s,m} = (C_{s,m} - MC_{s,m})/SD_{s,m}$$

where SD is standard deviation. Yearly standardized anomalies were created by averaging the monthly standard anomalies in a given year. Monthly climatologies demonstrate spatial and temporal changes while monthly standardized anomalies indicate where in space and time water quality metrics deviate from the annual average.

Satellite data sources

Satellite remote sensing data was collected by the Ocean Land Color Imager (OLCI) onboard the European Commission Copernicus programmes Sentinel-3A and -3B satellites. OLCI currently contains the most spectral bands and the highest spatial resolution (300 m) of the current suite of missions designed for ocean color and provides nearly daily coverage of Chesapeake Bay, making it an ideal sensor to study complex, small-scale dynamic processes typical of a coastal water body. Top-of-atmosphere full resolution (EFR), non-time critical Level-1 OLCI scenes overlapping the Chesapeake Bay region (36.8, −77.35, 39.65, −75.6) between April 2016 and October 2022 were obtained from various data portals (Copernicus Online Data Access REProcessed, EUMETSAT Data Center Long-Term Archive, and Copernicus

Online Data Access). The data were atmospherically corrected using the Case 2 Regional Coast Color (C2RCC) atmospheric correction algorithm which was demonstrated to be the best performing atmospheric correction method in Chesapeake Bay waters (Windle et al., 2022). Level-2 OLCI data were reprojected to the World Geodetic System 1984 coordinate system, merged when two scenes were collected on the same day, and clipped to a Chesapeake Bay shoreline shapefile. The final dataset included a total of 2,038 scenes including R_{rs} at 12 wavelengths (400, 412, 443, 490, 510, 560, 620, 665, 674, 681, 708 nm).

Optical water type classification

K-means clustering was employed to statistically partition R_{rs} to define fifteen distinct OWTs for Chesapeake Bay waters. To remove outliers from the clustering approach, pixels containing the highest 5% of each R_{rs} waveband were masked. To remove instances of potential land adjacency effect and shallow water, pixels within 1 km of the Chesapeake Bay shoreline were masked. Following methods from previous optical clustering studies, R_{rs} spectra were normalized (trapezoidal integration) to minimize the effect of variation in magnitude attributed to differences in optically active constituents and preserve the shape of R_{rs} across different parts of the spectrum (Vantrepotte et al., 2012; Spyrakos et al., 2018; Monolisha et al., 2018).

$$Rrs_{norm}(\lambda) = \frac{Rrs(\lambda)}{\int_{400}^{708} Rrs(\lambda) d\lambda}$$

A k-means clustering approach partitioned normalized R_{rs} (n = 64,386,450) into fifteen clusters of equal variance where each R_{rs} spectra belongs to the cluster with the nearest mean (cluster centers). Specifically, the sklearn.cluster.KMeans (scikit-learn v. 0.23.2) algorithm selected cluster centers that minimized the within-cluster sum of squares criterion:

$$\sum_{i=0}^{n} \min \left(\left| \left| x_i - u_i \right| \right|^2 \right)$$

Where n is a cluster, x_i is normalized R_{rs} spectra, and u_j is the mean normalized R_{rs} spectra in n cluster. The algorithm was run 10 times with different centroid means (n_init = 10) that were determined using random centroid initialization (random_state = 42) with a maximum of 300 iterations for a single run (max_iter = 300). The algorithm was run with a pre-selected fifteen clusters (n_clusters = 15) to characterize Chesapeake Bay OWTs. The algorithm was run on a range of pre-selected number of clusters, and the value of fifteen was selected to best represent the optical diversity of Chesapeake Bay and was close to the optimum number similar studies have used for coastal and inland waters (Melin & Vantrepotte, 2015 used 16; Spyrakos et al., 2018 used 9 and 12). The k-means clustering algorithm was run on each OLCI scene where each pixel was assigned an OWT according to its R_{rs} spectra. Individual OLCI scenes with corresponding OWT numbers were exported as new individual netCDF files. OWT data files were grouped by month and season and the count of each pixel was calculated to map dominant OWT occurrence across space and time.

Matchup analysis

A matchup analysis through space and time was conducted with all 41 mainstem and 24 tributary stations from the CBP Tidal Water Quality Monitoring database to analyze the underlying *in situ* data corresponding to each OWT. Water quality metrics that did not reveal a normal distribution were log transformed. *In situ* water quality constrained within each OWT were plotted as box and whisker plots within violin plots which show the smoothed kernel density estimation or the shape of the underlying distribution. The kernel bandwidth was calculated using Scott's rule of thumb (bw = 'scott') where the bandwidth is proportional to $n^{(-1(d+4))}$ where *n* is the number of points and *d* is the number of dimensions (Scott, 1992). A pairwise Tukey-HSD (honestly significant difference) post-hoc test was used to assess the significance of differences between pairs of group means for each water quality metric.

Algorithm selection and blending

For each OWT, seven existing empirical algorithms found in the literature were parameterized with corresponding *in situ* log transformed chlorophyll *a* and evaluated to determine the most effective model for retrieving chlorophyll *a* in the Chesapeake Bay. The following empirical chlorophyll *a* algorithms were evaluated:

Model A is NASA's OC4 band ratio algorithm which relates log transformed chlorophyll *a* to the maximum ratio of blue (443, 490, 510 nm) to green (560 nm) (O'Reiley et al., 2000);

$$\text{Chla\_A} = 10^{a + bX + cX^2 + dX^3 + eX^4}$$

where $X = \log 10(\text{Rrs}(443) > \text{Rrs}(490) > \text{Rrs}(510) / \text{Rrs}(560)$ and *a, b, c, d,* and *e* are NASA's OLCI coefficients that have been derived empirically from the NASA bio-Optical Marine Algorithm Dataset (NOMAD).[6]

Model B is a two-band ratio NIR-red algorithm employed in Dall'Olmo et al. (2003) and Moses et al. (2009). It is based on a linear relationship between *in situ* chlorophyll *a* and the ratio of R_{rs} in the NIR (708 nm) and red (665 nm) wavelengths;

$$\text{Chla\_B} = a \times \frac{\text{Rrs}(708)}{\text{Rrs}(665)} + b$$

where *a* and *b* are determined empirically for each OWT.

Model C is a three-band NIR-red algorithm that uses a third NIR band in an attempt to remove absorption due to NAP and CDOM which is assumed to be constant at both red wavelengths (Moses et al., 2009);

$$\text{Chla\_C} = a \times \left(\left[\frac{1}{\text{Rrs}(665)} - \frac{1}{\text{Rrs}(708)}\right] \times \text{Rrs}(753)\right) + b$$

where *a* and *b* are determined empirically for each OWT.

Model D is a two-band empirical ratio NIR-red algorithm employed by Gurlin et al. (2011);

$$\text{Chla\_D} = a \times \left(\frac{\text{Rrs}(708)}{\text{Rrs}(665)}\right)^2 + b \times \left(\frac{\text{Rrs}(708)}{\text{Rrs}(665)}\right) + c$$

where *a, b*, and *c* are determined empirically for each OWT.

---

[6] NASA Ocean Color Web. Chlorophyll a. Accessed online April 1, 2023: https://oceancolor.gsfc.nasa.gov/atbd/chlor_a/

Model E is a three-band empirical ratio algorithm employed by Gurlin et al. (2011);

$$\text{Chla\_E} = a \times \left(\left[\frac{1}{\text{Rrs}(665)} - \frac{1}{\text{Rrs}(708)}\right] \times \text{Rrs}(753)\right)^2 + b \times \left(\left[\frac{1}{\text{Rrs}(665)} - \frac{1}{\text{Rrs}(708)}\right] \times \text{Rrs}(753)\right) + c$$

where *a, b,* and *c* are determined empirically for each OWT.

Model F is the normalized difference chlorophyll index (NDCI) proposed by Mishra and Mishra (2012) which uses a two-band difference ratio;

$$\text{Chla\_F} = a + b \times \left(\frac{\text{Rrs}(708) - \text{Rrs}(665)}{\text{Rrs}(708) + \text{Rrs}(665}\right) + c \times \left(\frac{\text{Rrs}(708) - \text{Rrs}(665)}{\text{Rrs}(708) + \text{Rrs}(665}\right)^2$$

where *a, b*, and *c* are determined empirically for each OWT.

Model G refers to the NASA fluorescence line height (FLH) algorithm demonstrated by Gower et al. (1999) which estimates the magnitude of chlorophyll fluorescence at 681 nm above a baseline interpolated between 665 and 708 nm;

$$\text{FLH} = \text{Rrs}(681) - (\text{Rrs}(708) + (\text{Rrs}(665) - \text{Rrs}(708)) \times \frac{\lambda_{708} - \lambda_{681}}{\lambda_{708} - \lambda_{665}}$$

$$\text{Chla\_G} = a + b \times \text{FLH}$$

where *a* and *b* are determined empirically for each OWT.

Performance assessment of each chlorophyll *a* algorithm was measured by recommended statistical metrics for remote sensing data (Brewin et al., 2015; Seegers et al., 2018). The modeled versus observed chlorophyll *a* estimates were evaluated using a type II linear regression with the reduced major axis (RMA) approach to account for uncertainties in both the dependent and independent variables (Seegers et al., 2018). Regression slope, intercept, and the Pearson correlation coefficient (*r*) were used for performance assessment with a slope and *r* close to one and an intercept close to zero as an indication that the model compares well with the *in situ* data (Brewin et al., 2015). Additional performance metrics included mean absolute error (MAE) and root mean square error (RMSE) calculated as follows:

$$MAE = \frac{\sum_{i=1}^{n} |M_i - O_i|}{n}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (M_i - O_i)^2}{n}}$$

Where MAE indicates error magnitude between the paired modeled and observed chlorophyll *a* measurements and RMSE indicates variability in the error magnitudes (Seegers et al., 2018). Based on performance metrics, the best performing chlorophyll *a* algorithm for each OWT was selected to be applied in a dynamic switching algorithm to produce improved blended chlorophyll retrievals for the Chesapeake Bay.

Results

Monthly variability of in situ water quality data

Monthly climatologies and standardized anomalies of surface water quality data collected at 41 mainstem fixed stations from 2000 to 2022 is illustrated in Figure 2. In general, salinity follows a latitudinal gradient with lower values in the Upper

Bay and higher values in the Lower Bay. Seasonally, salinity is lowest in the spring and summer and maximal in the late fall, likely associated with freshwater flow and evaporative cooling, respectively. Overall, PN:PP values are higher than Redfield (16:1) throughout most of the Bay, with a few exceptions in the Upper Bay. PN:PP is consistently low across time in the Upper and Lower Bay. Across most stations, PN:PP is maximal in March and April, with the exception of the northernmost stations where PN:PP peaks in the summer. Surface chlorophyll *a* broadly mirrors PN:PP where the highest concentrations occur in the mesohaline Upper-Mid Bay. The timing of station specific chlorophyll *a* maxima varies, three stations had the highest monthly values in March, twelve in April, six in May, one in June, five in July, three in August, and three between September and December. Chl:C is also high in the spring throughout the Bay; however, some stations in the Mid to Lower Bay have maxima in the winter. TSS concentrations are highest in the Upper Bay with corresponding maxima in the spring, likely associated with higher freshwater flow. In the Mid and Lower Bay, the timing of monthly maxima is more varied, indicating variability in TSS concentrations. Organic and inorganic particle concentrations demonstrate opposing temporal patterns in the Mid to Lower Bay; with organic particle maxima occurring in the summer and inorganic particle maxima occurring in the fall and winter. Across time, Z_{SD}^{-1} and K_D show similar patterns with turbid waters in the Upper Bay and clear waters in the Lower Bay. Z_{SD}^{-1} and K_d maxima (most turbid water) occur in the spring in the Upper Bay but significantly differ in the Mid to Lower Bay with K_d being more variable across time.

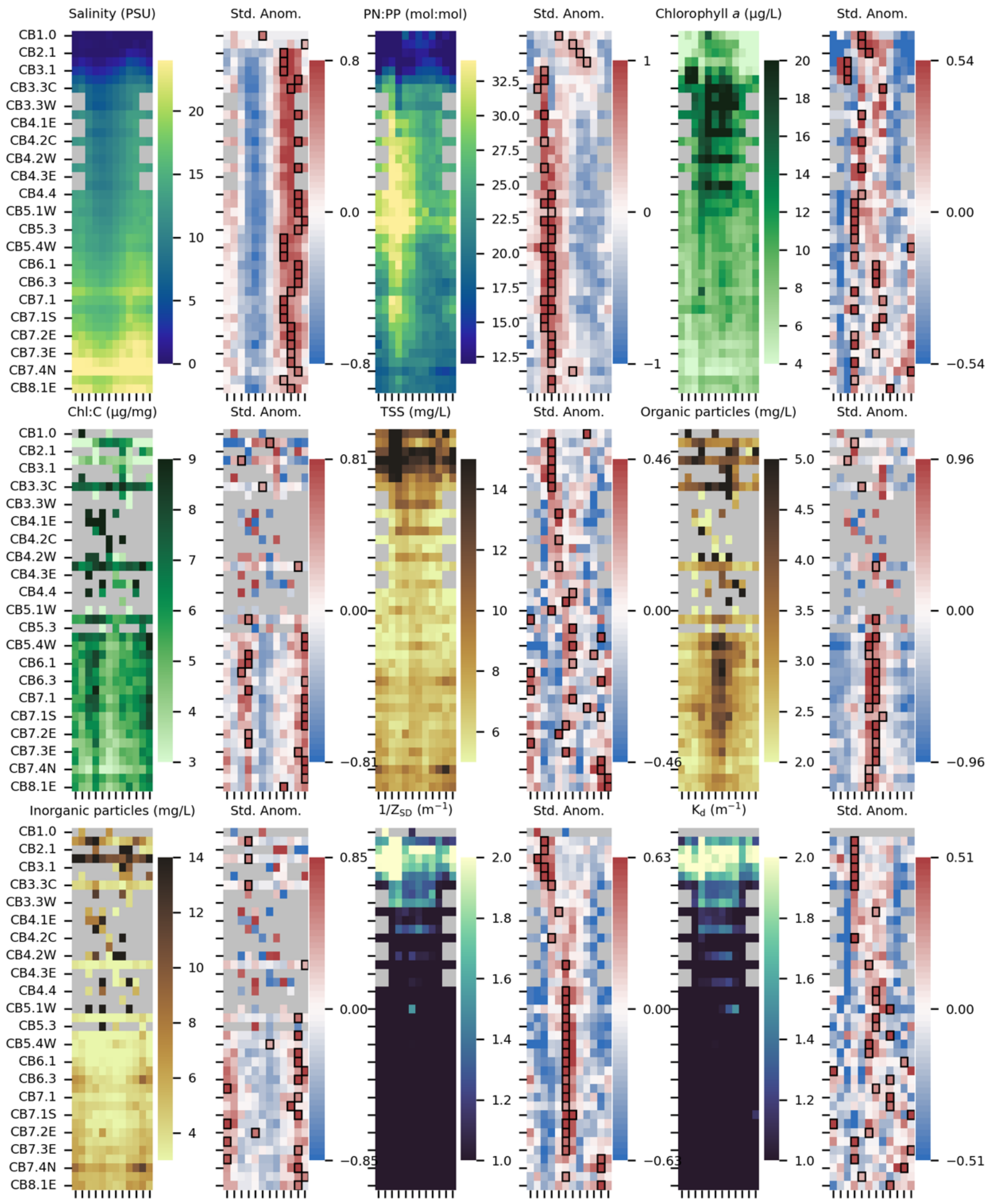


Figure 2. Monthly climatology and standardized anomaly plots of surface water quality variables collected at 41 fixed stations in the mainstem of the Chesapeake Bay between 2000 and 2022. Each

color bar is scaled to the 5th and 95th percentile. Maxima monthly standardized anomalies are outlined in black.

Interannual variability of in situ water quality

Interannual variability of surface water quality data collected at 41 mainstem fixed stations from 2000 to 2022 is represented as standardized anomalies alongside seasonal freshwater flow from the Susquehanna River (millions of $m^3 yr^{-1}$) in Figure 3. Between 2000 and 2022, annual freshwater flow from the Susquehanna River averaged 73,462 millions of $m^3 yr^{-1}$, was highest in 2018 with 115,643 $m^3 yr^{-1}$, and lowest in 2001 with 40,557 $m^3 yr^{-1}$. PN:PP was higher during years of low flow and lower during years of high flow. During some years of low flow, such as in 2001, PN:PP transitions with the Upper Bay containing higher values than the Lower Bay, illustrating a potential shift in nutrient limitation. Similarly, chlorophyll *a* was low during years of low flow and high during years of high flow. In general, chlorophyll *a* is higher in the Upper and Mid Bay compared to the Lower Bay, with some exceptions during high flow years when values are high in the Lower Bay. TSS portrays similar patterns to chlorophyll *a*, with higher concentrations during years of high flow. However, TSS concentrations are more variable, with some high flow years, such as 2018, having variable concentrations throughout the Bay. Organic suspended particles are variable with some years of high flow such as 2003 having high concentrations throughout the Bay and other years of high flow showing more variability, such as 2018. Suspended inorganic particle concentrations also show variability with freshwater flow. During years of high flow such as 2003, concentrations are low; however, in 2018, concentrations are high. During years of low flow such as in 2001, inorganic suspended particles are also high. Overall, inorganic suspended particle maxima seem to have increased in the last decade (2012-2022) while organic suspended particles have stayed variable. Z_{SD}^{-1} and K_d showed similar patterns with low values (clear water) during years of low flow (2002) contrasted with high values (turbid waters) during years of high flow (2011, 2018). However, K_d maxima was more variable across stations than Z_{SD}^{-1}. For example, in the high flow year of 2018, almost all stations contain high values of Z_{SD}^{-1}, while K_d is more variable. Both Z_{SD}^{-1} and K_d also seem to be more variable in the Lower Bay.

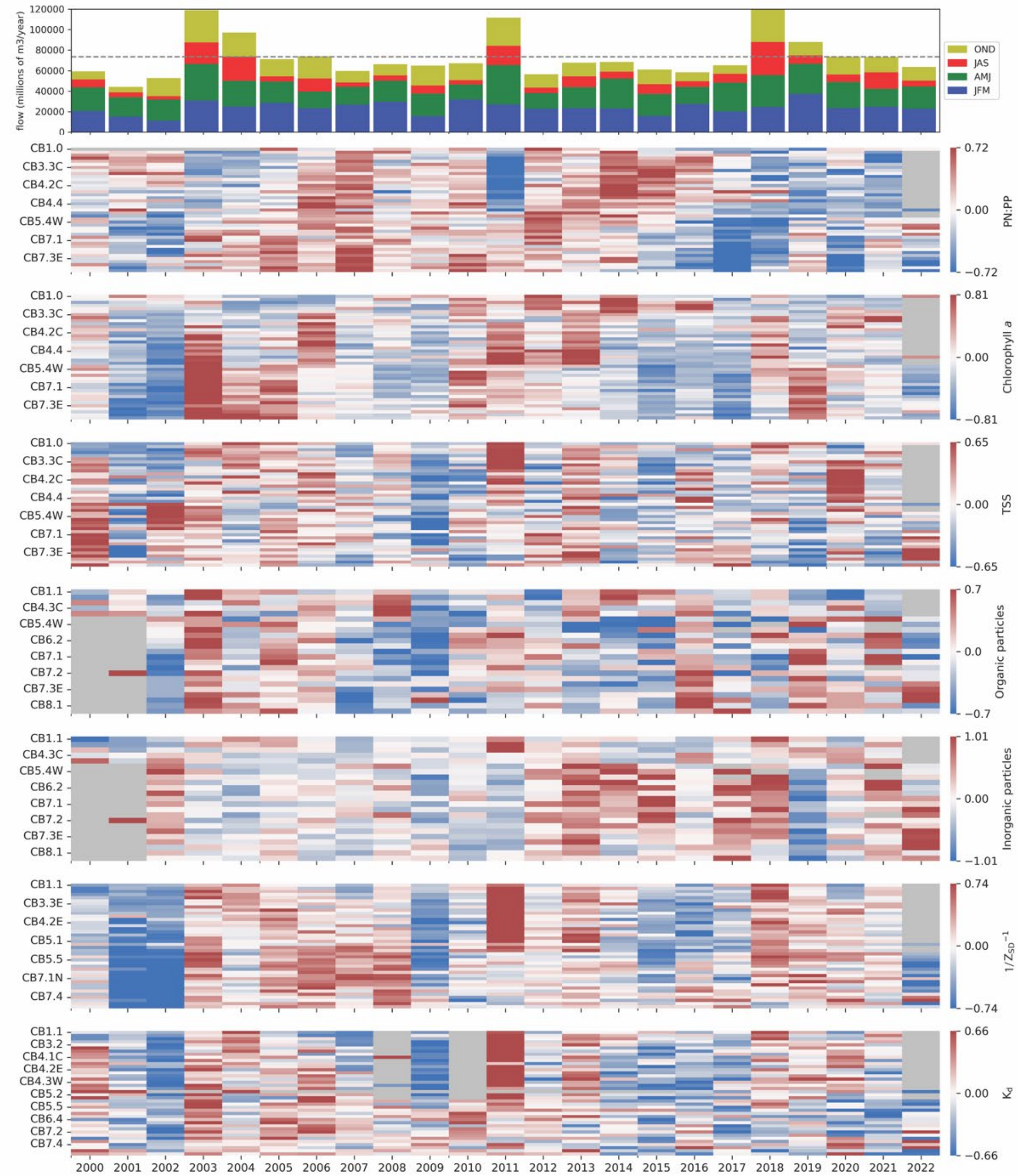


Figure 3. Annual Susquehanna River flow (millions of m^3/year) divided into seasonal contributions (top plot) and annual standardized anomaly plots of surface water quality parameters collected from 41 mainstem fixed stations between 2000 and 2022.

Optical water types in Chesapeake Bay

The *k*-means algorithm produced fifteen dominant normalized R_{rs} spectra, or OWTs, for the entire OLCI dataset (Figure 4). Most spectra peak in the green (560 nm), with OWT 7 peaking in the blue and OWT 6 and 8 peaking in the red. The percentage of OLCI pixels corresponding to each OWT ranged from 2 to 24% with OWT 14 containing the most and OWT 5 containing the least. The seasonal occurrence of each OWT is illustrated in Figure 5, note that the scale bar changes between OWTs. The matchup analysis with *in situ* water quality data is described in Table S1 and illustrated in Figure 6 which labels OWTs by the order of median chlorophyll *a*. The amount of *in situ* data constrained in each OWT ranged from 23 (OWT 7) to 387 (OWT 4). Results of the Tukey HSD test assessing the significance of differences between each OWT are summarized in Table S2.

OWT 1, representing 7% of data, commonly occurs in the Mid and Lower mainstem in the spring and summer. Amongst OWTs, these waters have the lowest median chlorophyll *a* concentration (5.62 ± 3.79 μg/L), fifth lowest median TSS concentration (7.00 ± 2.80 mg/L), third lowest median K_d (0.73 ± 0.49 m^{-1}), fourth lowest median TP concentration (0.03 ± 0.04 mg/L), and seventh lowest median TN concentration (0.59 ± 0.34 mg/L).

OWT 2, representing 7% of data, most commonly occurs in the mid Bay particularly near the Tangier Sound in the mid-Eastern Shore, and to a lesser degree the meso- and polyhaline waters of the mainstem in the winter and spring. Despite having the second lowest median chlorophyll *a* concentration (5.77 ± 4.30 μg/L), these are relatively turbid eutrophic waters as they correspond to the sixth highest median TSS concentration (8.50 ± 9.07 mg/L), seventh highest median TP concentration (0.04 ± 0.02 mg/L), and fourth highest median TN concentration (0.86 ± 0.36 mg/L) and K_d (1.48 ± 0.47 m^{-1}).

OWT 3, representing 4% of data, primarily occurs in the Upper Bay and upper tributaries in the summer. These are relatively transparent, low-nutrient waters with the third lowest median chlorophyll *a* concentration (6.65 ± 5.40 μg/L), second lowest median TN concentration (0.44 ± 0.01 mg/L) and median K_d (0.68 ± 0.35 m^{-1}), and the lowest median TP (0.02 ± 0.01 mg/L), and median TSS concentrations (5.20 ± 3.03 mg/L).

OWT 4, representing 4% of data, primarily occurs in the lower Mid Bay in the spring and along the lower Eastern Shore in the summer. OWT 4 represents the most transparent waters with the fourth lowest median chlorophyll *a* concentration (6.88 ± 7.28 μg/L), third lowest median TP (0.03 ± 0.01 mg/L), and TSS concentrations (6.00 ± 1.96 mg/L), and the lowest TN (0.39 ± 0.19 mg/L) and K_d (0.67 ± 0.37 m^{-1}).

OWT 5, representing 1% of data, most commonly occurs in the Mid and Lower Bay mainstem in the summer. It depicts waters with the fifth lowest median chlorophyll *a* (7.37 ± 4.31 μg/L) and median TP concentrations (0.03 ± 0.02 mg/L), and the fourth lowest median TN (0.48 ± 0.31 mg/L), TSS concentrations (6.00 ± 11.26 mg/L), and K_d (0.74 ± 0.24 m^{-1}).

OWT 6, representing 5% of data, is one of two spectra to peak in the red and most commonly occurs in the upper Mid Bay in the summer and fall. These are the most turbid waters with the sixth lowest median chlorophyll *a* concentration (7.48 ± 7.44 μg/L), third highest median TP (0.05 ± 0.02 mg/L), second highest median TN (1.01 ± 0.45 mg/L), and the highest median TSS concentrations (16.0 ± 11.7 mg/L) and K_d (2.27 ± 0.33 m^{-1}).

OWT 7, representing 6% of data, depicts the bluest of Chesapeake Bay waters and most commonly occurs in the Mid and Low mainstem, particularly in summer. OWT 7 has the seventh lowest median chlorophyll *a* (5.59 ± 2.70 μg/L), the fifth lowest median TN concentration (0.55 ± 0.20 mg/L) and K_d (0.77 ± 0.33 m^{-1}), and the second lowest TP (0.03 ± 0.02 mg/L) and TSS concentrations (5.59 ±2.70 mg/L).

OWT 8, representing 4% of data, is one of two spectra to peak in the red and primarily occurs in the tidal fresh Upper Bay and tributaries throughout the year. It depicts turbid waters with the eighth highest median chlorophyll *a* concentration (8.54 ± 20.2 μg/L), the second highest median TP (0.05 ± 0.02 mg/L), second highest median TSS (14.4 ± 16.7 mg/L), second highest median K_d (1.92 ± 0.58 m^{-1}), and the highest median TN concentrations (1.03 ± 0.39 mg/L).

OWT 9, representing 4% of data, most commonly occurs in the Upper Bay and tributaries in the fall and winter. OWT 9 depicts waters with the ninth highest median chlorophyll *a* (8.89 ± 4.00 μg/L), eighth highest median TP (0.03 ± 0.01 mg/L), fifth lowest median TN (0.58 ± 0.29 mg/L), seventh lowest median TSS (7.10 ± 6.96 mg/L), and median K_d (0.91 ± 0.37 m^{-1}).

OWT 10, representing 2% of data, occurs primarily in the tidal fresh Upper Bay in the fall. OWT 10 depicts relatively turbid waters with the sixth highest median chlorophyll *a* (8.94 ± 6.14 μg/L), sixth lowest median TP (0.03 ± 0.01 mg/L), third lowest median TN (0.46 ± 0.19), eighth lowest highest median TSS (7.13 ± 3.13 mg/L), and sixth lowest median K_d (0.79 ± 0.32 m^{-1}).

OWT 11, representing 11% of data, most commonly occurs in the Mid and Lower Bay mainstem in the fall and winter. OWT 11 depicts moderately turbid waters with the fifth highest median chlorophyll *a* (9.19 ±7.49 μg/L), median TP (0.04 ± 0.01 mg/L), and median TN (0.78 ± 0.40 mg/L), and the fourth highest median TSS (10.3 ± 9.18 mg/L) and K_d (1.42 ± 0.32 m^{-1}).

OWT 12, representing 4% of data, represents relatively turbid waters and is located primarily in the tidal fresh Upper Bay throughout time. It depicts waters with the fourth highest median chlorophyll *a* (9.77 ± 6.22 μg/L), sixth highest median TN (0.66 ± 0.29 mg/L), seventh highest median TP (0.04 ± 0.02 mg/L), fourth highest median TSS (9.60 ± 8.63 mg/L), and fifth highest median K_d (1.41 ± 0.45 m^{-1}).

OWT 13, representing 4% of data, predominantly occurs in the upper Mid Bay and middle reaches of the Potomac River in the winter and spring and represents waters with the third highest median chlorophyll *a* (10.7 ± 5.99 μg/L), seventh lowest median TP (0.03 ± 0.01 mg/L), eighth lowest median TN (0.63 ± 0.21 mg/L), and K_d (0.98 ± 0.41 m^{-1}) and fifth lowest median TSS (6.90 ± 3.83 mg/L).

OWT 14, the most abundant OWT representing 24% of data, most commonly occurs in the Lower Bay mainstem in the summer. OWT 14 represents relatively turbid waters with the second highest median chlorophyll *a* (15.0 ± 47.2 μg/L), highest median TP (0.05 ± 0.02 mg/L), and the third highest median TN (0.86 ± 0.74 mg/L), TSS (11.2 ± 14.5 mg/L), and K_d (1.52 ± 0.68 m^{-1}).

OWT 15, representing 13% of data, most commonly occurs in the Mid and Lower Bay mainstem in the summer and fall. It depicts waters with the highest median chlorophyll *a* concentration (16.3 ± 11.7 μg/L), the fourth highest median TP (0.04 ± 0.02 mg/L), sixth highest median TN (0.71 ± 0.21 mg/L), seventh highest median TSS (8.00 ± 3.65 mg/L) and median K_d (1.29 ± 0.36 m^{-1}).

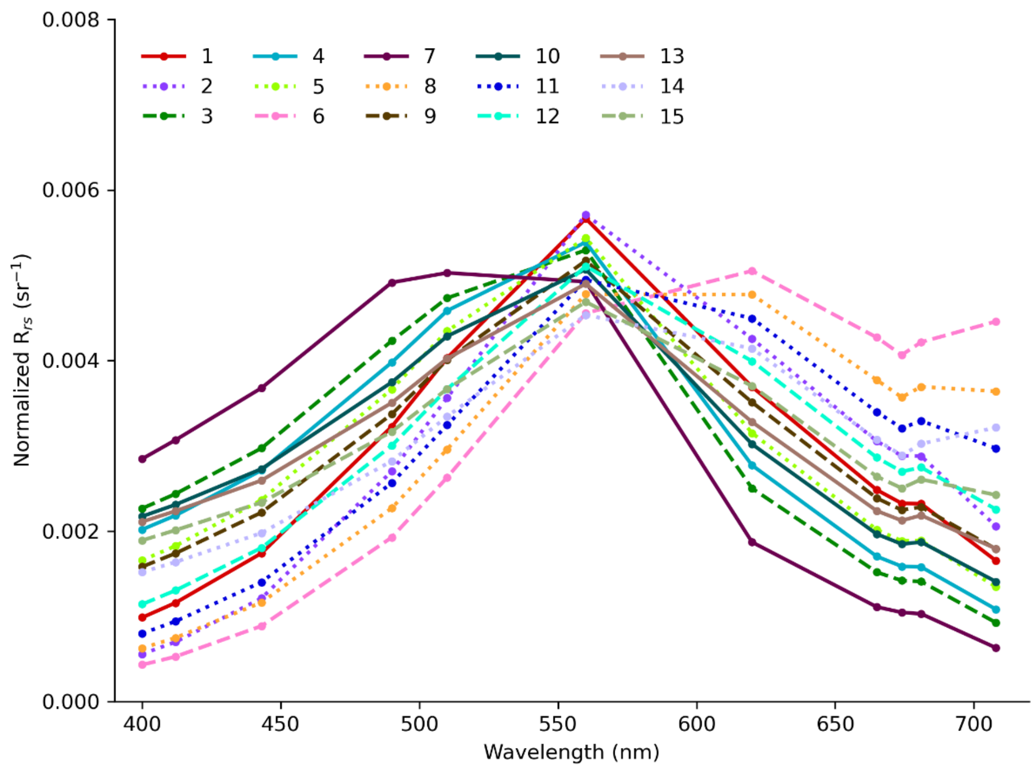
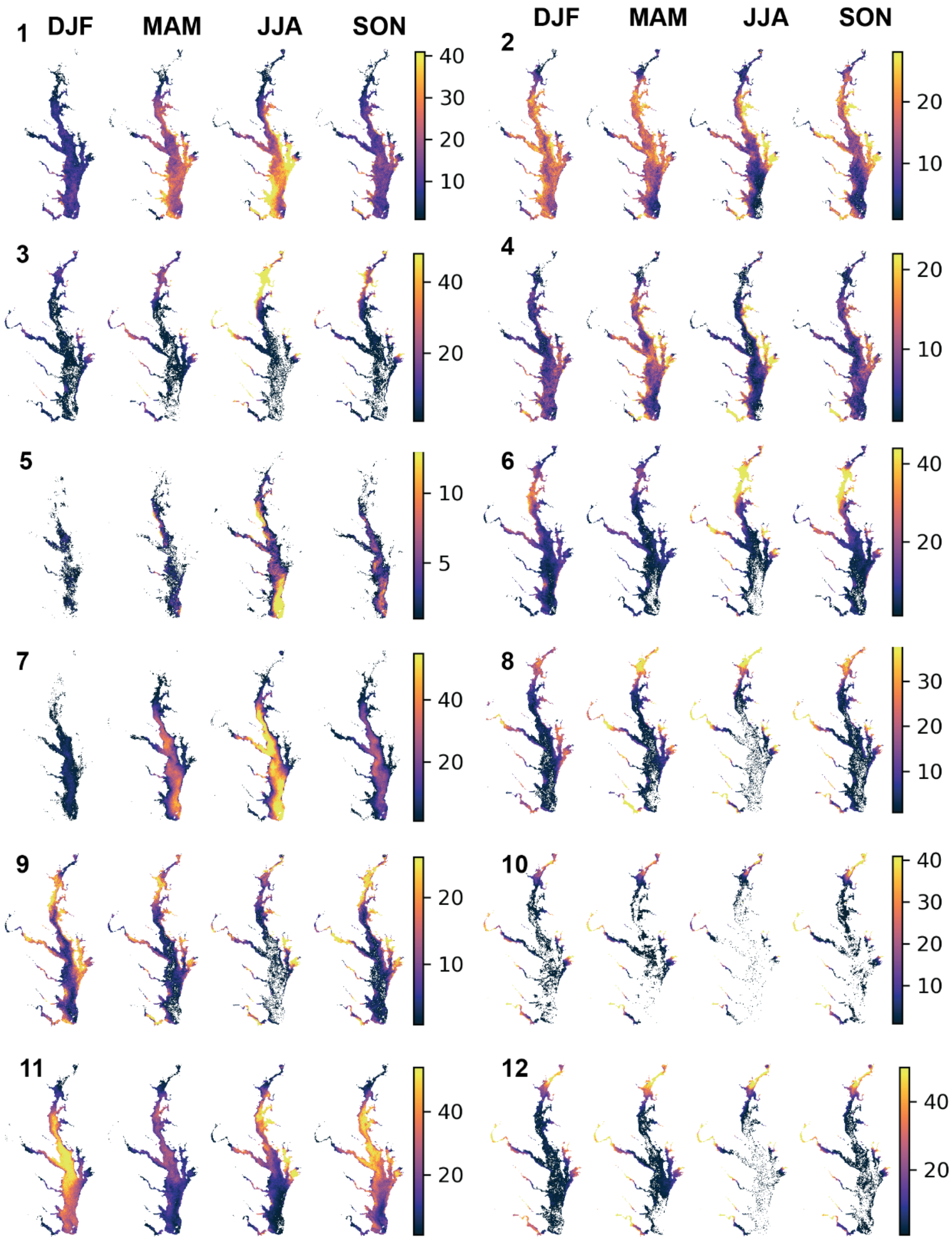


Figure 4. Normalized R_{rs} spectra for each OWT in Chesapeake Bay as identified by *k*-means clustering algorithm.


1
DJF
MAM
JJA
SON
2
DJF
MAM
JJA
SON
3
4
5
6
7
8
9
10
11
12

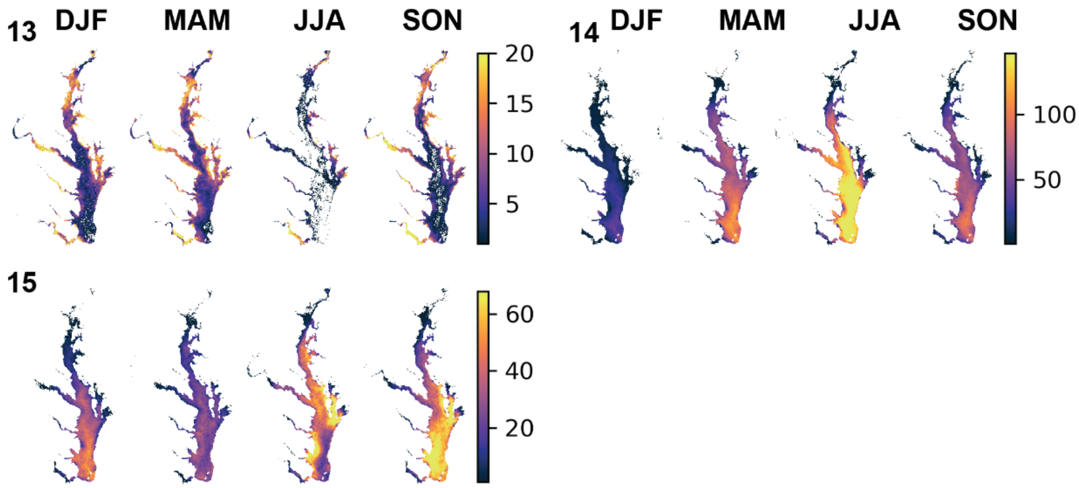


Figure 5. The average seasonal occurrence of each OWT. Note that the scale bar changes between OWTs.

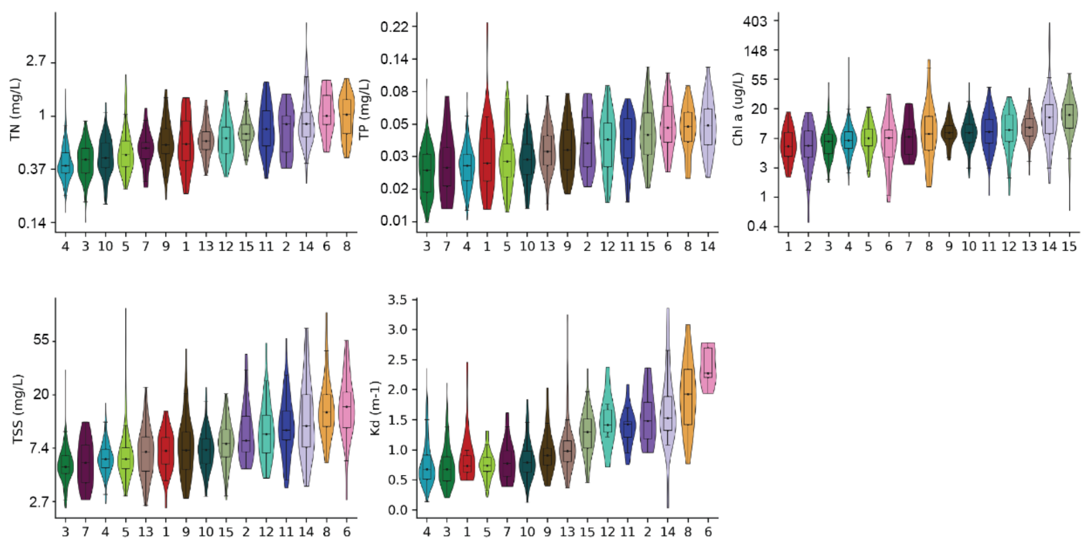


Figure 6. Boxplots with probability density of A) TN (mg/L), B) TP (mg/L), C) chlorophyll *a* (µg/L), D) TSS (mg/L), and E) K_d (m^{-1}) for each optical water type. The sample interquartile range (25, 50, 75%) is indicated by the horizontal lines within the box while tails represent maximum and minimum values. OWTs are labeled in the order of median chlorophyll *a*.

The monthly occurrence of each OWT is illustrated in Figure 7, with OWTs colored by lowest to highest K_d. In general, OWTs representing higher K_d values (lighter colors) have higher percent occurence during times of high flow whereas OWTs representing lower K_d values (darker colors) have lower percent occurrence during times of low flow.

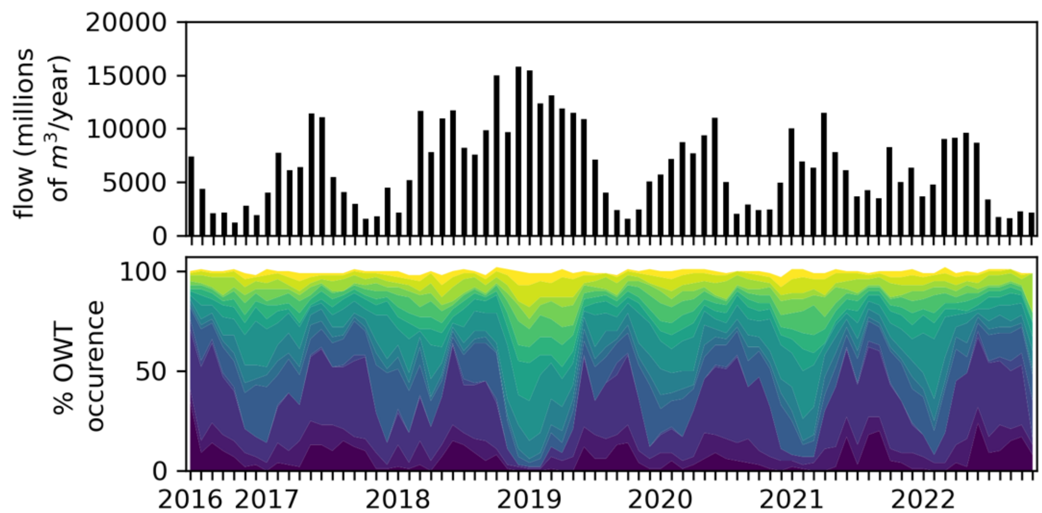


Figure 7. Monthly Susquehanna River flow (millions of m^3/year) (top plot) and a stacked plot showing the percent occurrence of OWTs within each month, with each color representing an OWT ordered from low (dark colors) to high K_d (lighter colors) (bottom plot).

Selection of and blending chlorophyll a algorithms

Performance metrics for each chlorophyll *a* algorithm and OWT are summarized in Table S3, and the best algorithm selected for each OWT is summarized in Table 1. Overall, Model F performed best for five OWTs (4, 5, 6, 8, 11), Model C performed best for four OWTs (2, 3, 7, 10). Model E performed best for three OWTs (1, 9, 14), Model D performed best for two OWTs (12, 15), and Model G performed best for OWT 14. Model A and B were never selected as the best performing algorithm.

Table 1. Recommended empirical chlorophyll *a* algorithm and coefficients calibrated for each OWT.

OWT	Model	Equation	*a*	*b*	*c*
1	E	$\text{Chla\_E} = a \times \left(\left[\frac{1}{\text{Rrs}(665)} - \frac{1}{\text{Rrs}(708)}\right] \times \text{Rrs}(753)\right)^2 + b \times \left(\left[\frac{1}{\text{Rrs}(665)} - \frac{1}{\text{Rrs}(708)}\right] \times \text{Rrs}(753)\right) + c$	698.68	128.33	6.62
2	C	$\text{Chla\_C} = a \times \left(\left[\frac{1}{\text{Rrs}(665)} - \frac{1}{\text{Rrs}(708)}\right] \times \text{Rrs}(753)\right) + b$	0.00	0.35	
3	C	$\text{Chla\_C} = a \times \left(\left[\frac{1}{\text{Rrs}(665)} - \frac{1}{\text{Rrs}(708)}\right] \times \text{Rrs}(753)\right) + b$	0.00	0.62	
4	F	$\text{Chla\_F} = a + b \times \left(\frac{\text{Rrs}(708) - \text{Rrs}(665)}{\text{Rrs}(708) + \text{Rrs}(665}\right) + c \times \left(\frac{\text{Rrs}(708) - \text{Rrs}(665)}{\text{Rrs}(708) + \text{Rrs}(665}\right)^2$	-0.48	-15.61	-43.01
5	F	$\text{Chla\_F} = a + b \times \left(\frac{\text{Rrs}(708) - \text{Rrs}(665)}{\text{Rrs}(708) + \text{Rrs}(665}\right) + c \times \left(\frac{\text{Rrs}(708) - \text{Rrs}(665)}{\text{Rrs}(708) + \text{Rrs}(665}\right)^2$	2.34	11.16	18.81
6	F	$\text{Chla\_F} = a + b \times \left(\frac{\text{Rrs}(708) - \text{Rrs}(665)}{\text{Rrs}(708) + \text{Rrs}(665}\right) + c \times \left(\frac{\text{Rrs}(708) - \text{Rrs}(665)}{\text{Rrs}(708) + \text{Rrs}(665}\right)^2$	0.69	0.20	32.17
7	C	$\text{Chla\_C} = a \times \left(\left[\frac{1}{\text{Rrs}(665)} - \frac{1}{\text{Rrs}(708)}\right] \times \text{Rrs}(753)\right) + b$	0.00	0.66	
8	F	$\text{Chla\_F} = a + b \times \left(\frac{\text{Rrs}(708) - \text{Rrs}(665)}{\text{Rrs}(708) + \text{Rrs}(665}\right) + c \times \left(\frac{\text{Rrs}(708) - \text{Rrs}(665)}{\text{Rrs}(708) + \text{Rrs}(665}\right)^2$	0.87	2.99	25.18
9	E	$\text{Chla\_E} = a \times \left(\left[\frac{1}{\text{Rrs}(665)} - \frac{1}{\text{Rrs}(708)}\right] \times \text{Rrs}(753)\right)^2 + b \times \left(\left[\frac{1}{\text{Rrs}(665)} - \frac{1}{\text{Rrs}(708)}\right] \times \text{Rrs}(753)\right) + c$	-12.79	3.76	1.31
10	C	$\text{Chla\_C} = a \times \left(\left[\frac{1}{\text{Rrs}(665)} - \frac{1}{\text{Rrs}(708)}\right] \times \text{Rrs}(753)\right) + b$	0.00	0.84	

11	F	$\mathrm{Chla\_F} = a + b \times \left(\frac{\mathrm{Rrs}(708) - \mathrm{Rrs}(665)}{\mathrm{Rrs}(708) + \mathrm{Rrs}(665}\right) + c \times \left(\frac{\mathrm{Rrs}(708) - \mathrm{Rrs}(665)}{\mathrm{Rrs}(708) + \mathrm{Rrs}(665}\right)^2$	1.08	1.05	-6.26
12	D	$\mathrm{Chla\_D} = a \times \left(\frac{\mathrm{Rrs}(708)}{\mathrm{Rrs}(665)}\right)^2 + b \times \left(\frac{\mathrm{Rrs}(708)}{\mathrm{Rrs}(665)}\right) + c$	-8.36	15.16	-5.79
13	G	$\mathrm{Chla\_G} = a + b \times \mathrm{FLH}$ $\mathrm{FLH} = \mathrm{Rrs}(681) - (\mathrm{Rrs}(708) + (\mathrm{Rrs}(665) - \mathrm{Rrs}(708)) \times \frac{\lambda_{708} - \lambda_{681}}{\lambda_{708} - \lambda_{665}}$	1.46	-2749.42	
14	E	$\mathrm{Chla\_E} = a \times \left(\left[\frac{1}{\mathrm{Rrs}(665)} - \frac{1}{\mathrm{Rrs}(708)}\right] \times \mathrm{Rrs}(753)\right)^2 + b \times \left(\left[\frac{1}{\mathrm{Rrs}(665)} - \frac{1}{\mathrm{Rrs}(708)}\right] \times \mathrm{Rrs}(753)\right) + c$	19.00	-0.62	1.09
15	D	$\mathrm{Chla\_D} = a \times \left(\frac{\mathrm{Rrs}(708)}{\mathrm{Rrs}(665)}\right)^2 + b \times \left(\frac{\mathrm{Rrs}(708)}{\mathrm{Rrs}(665)}\right) + c$	-2.41	6.11	-2.39

Discussion

The Chesapeake Bay is one of the most well-studied estuarine systems in the world with a wealth of long-term water quality monitoring data. Surface water quality parameters collected from all mainstem fixed stations between 2000 and 2022 were synthesized to provide insight into the seasonal and interannual variability of water quality dynamics. To increase sampling coverage, satellite remote sensing data collected in Chesapeake Bay from 2016 to 2022 was also analyzed. This study is the first attempt to statistically classify OWTs from satellite remote sensing data collected in the Chesapeake Bay. OWTs were evaluated against *in situ* water quality data to characterize how each OWT corresponds to the underlying water quality data. This information was used to optimize the selection of bio-optical algorithms to produce a blended chlorophyll *a* product. Results have broad implications for improving satellite water quality retrievals in Chesapeake Bay and can be incorporated into water quality monitoring and management.

Seasonal variability of in situ water quality data

Water quality dynamics in coastal waters differ from the open ocean due to influences from the surrounding land, atmosphere, and underlying sediments that fluctuate over time (Cloern, 1996; Cloern & Jassby, 2010; Kemp et al., 2005). In general, the Bay experiences feedback mechanisms stemming from increased nutrient loading that can enhance phytoplankton production, deplete bottom water oxygen, increase nutrient cycling from sediments, and intensify eutrophication (Kemp et al., 2005). These relationships significantly vary due to regional- and local-scale processes from freshwater flow, wind-driven resuspension, tidal mixing and other physical mechanisms, and enriched nutrient inputs from proximity to land (Kemp et al., 2005; Roman et al., 2005; Cloern & Jassby, 2008). In this section, existing knowledge of Chesapeake Bay water quality is synthesized with seasonal variability of *in situ* observations of salinity, suspended particles, nutrients, chlorophyll *a*, and water clarity.

Salinity

The mouth of the Chesapeake Bay receives oceanic inputs from the Atlantic Ocean while the head of the Bay receives freshwater from rivers contributing to a partially mixed estuary with a latitudinal salinity gradient (Pritchard, 1952; Fisher et al., 1988; Boicourt, 1999). Salinity can be influenced by freshwater variability and tidal mixing, potentially contributing to some stations in the Mid and Lower Bay diverging from the latitudinal salinity gradient (Figure 2). In the summer and fall, salinity is high due to lower freshwater flow, higher rates of evaporation, and a deeper mixed layer that entrains high salinity waters to the surface. In the spring, salinity is low due to increased mixing from the spring freshet, an annual high flow event resulting from snow and ice melting (Goodrich et al, 1987; Schubel & Pritchard, 1986).

Suspended particles

Suspended particles, often referred to as total suspended sediments/solids (TSS) or total suspended matter (TSM), are a combination of inorganic sediments and allochthonous and autochthonous organic matter. The type and size of suspended particles in the Bay varies due to river discharge, regional geology, geomorphology, land uses, estuarine circulation, physical, and biological processes (Gellis et al., 2000; Langland & Cronin, 2003; Roman et al., 2005; Brakebill et al., 2010; Noe et al., 2020). The Bay is believed to be mostly composed of fine-grained inorganic particles since the three largest tributaries that export 90% of suspended particles flow through the Piedmont physiographic province which is characterized with fine-grained inorganic sediment (Zhang & Blomquist, 2018). Higher concentrations of inorganic particles can also be attributed to the dimished trapping capacity of the Conowingo Dam located on the Susquehanna River over the last twenty years (Zhang et al., 2016; Zhang & Blomquist, 2018). An abundance of suspended particles can lead to ecological consequences such as blocking light penetration for submerged aquatic vegetation, carrying toxic contaminants, and burying benthic habitats such as oyster beds (Kemp et al., 2005). Nutrients can also be transported via suspended particles; Zhang et al. (2015) estimated that 73% of the total phosphorus load and 18% of the total nitrogen load transported throughout the Bay is attached to inorganic suspended particles.

Surface TSS concentrations are high at Upper Bay stations across time due to consistent riverine inputs, particularly the Susquehanna River (Figure 2). In general, TSS maxima occurs in the spring; however, there is variability across space demonstrating that TSS is likely influenced by annual freshwater flow or other physical mechanisms. Maximal TSS values in the Upper Bay likely demarcate the estuarine turbidity maximum (ETM), a zone containing increased suspended particulate concentration associated with the landward limit of salt intrusion, known to trap settling particles that are resuspended by tidal currents (Sanford et al., 2001; Roman et al., 2001; North & Houde, 2003). Other physical discontinuities that tend to aggregate suspended particles could explain the variability in TSS maxima in the Mid Bay (Roman et al., 2005). Organic suspended particle concentration maxima occur mostly in the summer in the Mid and Lower Bay, potentially associated with an increase in phytoplankton biomass. The temporal mismatch with chlorophyll *a* maxima is consistent with a photoacclimative response in the summertime phytoplankton community to higher light and more prevalent nutrient limitation as both cause a down-regulation in pigmentation and associated photosynthetic machinery (Figure 2, Malone et al., 1998, Bronk et al., 1998). Studies have also observed higher organic particles in the Mid Bay in contrast with inorganic particles attributed to more shoreline armoring by human development (Russ & Palinkas, 2020; Turner et al., 2021). In contrast, inorganic suspended particle concentration maxima occur mostly in the winter in the Mid and Lower Bay, potentially due to resuspension from wind mixing and other physical mechanisms known to aggregate suspended particles (Roman et al., 2005). This summertime region of low suspended inorganic particles and high suspended organic particles could also be due to a shift in particle composition, as suggested in previous studies (Gallegos et al., 2011; Testa et al., 2019; Turner et al., 2021).

PN:PP

In an estuary, the major limiting nutrients to phytoplankton growth are nitrogen, phosphorus, and silica with the amount of each affecting phytoplankton uptake and growth (Fisher et al., 1992). The amount of these limiting nutrients impacts algal production and biomass, which has subsequent effects on higher trophic levels and water clarity (Kemp et al., 2005). Nitrogen and phosphorus enter the water via point or nonpoint sources with the latter being the dominant source of nutrient loading in the Bay, varying with meteorological and climatic conditions (Kemp et al., 2005). Nutrients are also recycled in the sediments and water column, delivering regenerated forms through diffusion and mixing (Malone et al., 1988; Boynton et al., 1995). Changes in the stoichiometric PN:PP ratio can provide insight into nutrient limitation in a water body (Redfield, 1963; Guildford & Hecky, 2000). Upper Bay stations contained consistently low PN:PP values and high TSS concentrations throughout the year, suggesting light limitation where phytoplankton growth is controlled by insufficient light, precluding growth responses to nutrient addition (Figure 2, Kemp et al., 2005). Mid Bay stations contain high PN:PP values in the winter and spring due to increased nitrogen loading and nutrient recycling, suggesting P limitation and shifts to lower values in the summer and fall likely due to reduced nitrogen loading via freshwater, high salinity, and high rates of denitrification and phosphate release from anoxic sediments (Boynton et al., 1982; Fisher et al., 1992; Malone et al. 1996; Kemp et al., 2005; Guildford & Hecky, 2010). Some stations in the Mid Bay demonstrate high PN:PP values throughout the year, likely due to close proximity of nitrogen loading from riverine inputs (Figure 2). Lower Bay stations contain consistently low PN:PP values throughout the year, suggesting nitrogen limitation (Figure 2, Harding et al., 1994; Xu and Hood). The PN:PP climatology demonstrated similar patterns to existing nutrient and light limitation conceptual models (Figure 2, Kemp et al., 2005; Zhang et al., 2021).

Chlorophyll a

Chlorophyll *a* is the main photosynthetic pigment used by phytoplankton to capture and convert light energy into biomass and can be used to study how phytoplankton productivity relates to seasonal changes in light, nutrients, and grazing (Malone et al., 1988; Harding & Perry, 1997; Kemp et al., 2005). Previous studies of the Bay observe chlorophyll *a* maxima in the spring when grazing rates are low and nutrient loads are high from the spring freshet (Malone et al., 1998; Bronk et al., 1998). As the Bay transitions to summer, temperatures increase, bottom organic matter gets remineralized, and phytoplankton growth rates increase, leading to a balance of growth and grazing and a maximum in phytoplankton productivity (Malone et al., 1991; Malone et al., 1996, Bronk et al., 1998; Xu & Hood, 2006). In the fall and winter, nutrient inputs are reduced, water becomes more mixed, and the average light exposed to phytoplankton gets lower, leading to lower phytoplankton biomass and chlorophyll *a* concentrations (Xu & Hood, 2006).

The magnitude, timing, and location of the surface chlorophyll *a* maximum can vary widely due to climatic variability, nutrient distributions, and physical processes (Harding et al., 1994; Roman et al., 2005; Adolf et al., 2006; Cloern & Jassby, 2010). The majority of stations contain high chlorophyll *a* concentrations in the spring likely associated with increased nutrient loading with the spring freshet; however, there is notable variability with some Upper and Mid Bay stations peaking in the summer and some Lower Bay stations peaking in the winter. This could be due to known physical mechanisms such as convergent circulation in the Lower Bay (Roman et al., 2005; Xu & Hood, 2006).

Chlorophyll a to carbon (Chl:C)

The amount of chlorophyll *a* measured is not always an accurate approximation of phytoplankton biomass. Phytoplankton continuously adjust their concentrations of chlorophyll *a* per cell, relative to that of carbon, in response to the magnitude of incident light and nutrient availability (Geider et al., 1996; Behrenfeld et al., 2005). Therefore, the ratio of Chl:C is a useful metric to infer phytoplankton photoacclimation and physiological state, thereby providing insight into growth rates and phytoplankton productivity (Behrenfeld et al., 2005; Sathyendranath et al., 2020). In general, Chl:C increases with light limitation and decreases with nutrient limitation. Chl:C is high at some Upper Bay stations with maximal values in the spring and summer, likely due to light limitation from high concentrations of inorganic particles (Figure 2). Mid and Lower Bay stations demonstrate high Chl:C both in the spring and winter (Figure 2) relative to the summer and fall, suggesting stronger nutrient limitation in the latter two seasons.

Water clarity

Water clarity refers to the distance that light penetrates through water, as well as the visibility of objects through water, and is influenced by the amount and type of suspended particles (Turner et al., 2022). Water clarity is an important metric to monitor since it has consequences for primary production, benthic habitats, harmful algal blooms, and overall ecosystem health (Kemp et al., 2005). Water clarity is typically measured using two methods: 1) measuring vertical profiles of downwelling irradiance, or photosynthetically active radiation (PAR), and estimating a coefficient of light attenuation (K_d, m^{-1}), and 2) measuring the Secchi disk depth (Z_{SD}, m), or how far a black and white disk can be seen by the human eye (Preisendorfer, 1986). Historically, these two measurements were thought to be highly correlated; however, recent studies have shown that the relationship between K_d and Z_{SD} can differ, and the mismatch can infer useful information about water quality constituents that limit light penetration (Gallegos et al., 2011; Turner et al., 2022). Smaller inorganic particles tend to scatter light in all directions while larger organic particles typically scatter more light in the forward (downward) direction (Stramski et al., 2001; Babin et al., 2003). Since a Secchi disc is known to be more sensitive to forward light scattering than K_d, changes in particle composition can lead to opposing trends of Z_{SD} and K_d (Hou et al., 2007; Gallegos et al., 2011). Gallegos et al. (2011) analyzed long-term (1985 to 2010) water clarity trends in the Mid Bay and found a slight decline in Z_{SD}, but stable K_d which led to a hypothesis that suspended particle concentrations in the Mid Bay have shifted from inorganic to organic particles, a region that Turner et al. (2021) has termed the 'Organic fog zone'. Changing suspended particle composition could also result from shifting phytoplankton community composition from larger to smaller cells, particularly in the Mid and Lower Bay where phytoplankton are more nutrient limited (Harding et al., 2015).

Patterns of Z_{SD}^{-1} and K_d are similar throughout the Chesapeake Bay with higher (less clear) values in the Upper Bay and a maximum in the spring likely due to high riverine inputs, phytoplankton biomass, and resuspension of particles (Figure 2, Testa et al., 2019). However, in the Mid and Lower Bay, the timing of maxima differs with Z_{SD}^{-1} being consistently higher in the summer, likely associated with high phytoplankton productivity and an increase in organic particles and K_d being more variable (Figure 2). This finding matches observations from previous studies of a shift in particle composition and therefore varying water clarity measurements in this region of the Bay (Gallegos et al., 2011; Turner et al., 2021).

Interannual variability

Interannual variability of meteorological and climatic conditions have significant impacts on Chesapeake Bay water quality and can complicate the detection of long-term trends (Harding et al., 2016; Testa et al., 2019). Nutrient loads in the form of runoff can vary by a factor of two between high and low flow years, impacting phytoplankton productivity and water clarity (Hagy et al., 2004; Harding et al., 2016). Similarly, higher freshwater flow can considerably increase TSS concentrations, influencing light penetration and phytoplankton growth (Harding et al., 1994). In this study, annual standardized anomalies of water quality parameters show marked differences in low and high flow years, demonstrating the need to adjust for freshwater flow to to better analyze where management actions are working or needed (Harding et al., 2016; Zhang & Blomquist, 2018; Murphy et al., 2022).

Between the years 2000 and 2022, the Bay received the least amount of freshwater in 2001 and the most in 2018 (Figure 3). In 2001, PN:PP is high in the Upper Bay, suggesting phosphorus limitation and low in the Lower Bay, suggesting nitrogen limitation; however, in 2018, PN:PP is low throughout the Bay, likely associated with high phosphorus loading through runoff (Figure 3). Chlorophyll *a* concentrations seem to correspond to PN:PP trends, with lower values in 2001 and higher values in 2018. In 2018, chlorophyll *a* maxima occur slightly further south, likely due to light limitation in the Upper Bay from increased suspended particles that block light and phytoplankton growth (Figure 3, Harding et al., 1994; Xu & Hood, 2006). TSS concentrations also correspond with flow, with overall lower values in 2001 and higher values in 2018. In 2018, suspended organic particle maxima are variable while suspended inorganic particle maxima are high, particularly in the Mid Bay, demonstrating that inorganic particles dominate during years of high flow. However, in other high flow years such as 2003, organic particles maxima are higher than inorganic particles, indicating particle composition is complex and likely influenced by multiple factors. In general, Z_{SD}^{-1} and K_d are similar with lower values (clearer water) in 2001 than in 2018 corresponding to freshwater flow. In general, K_d is more variable potentially due to shifting suspended particle composition (Figure 3, Gallegos et al., 2011; Turner et al., 2021).

OWT classification

The variability in the shape of the OWT spectra represents the underlying optical variability of the Chesapeake Bay (Figure 4). In general, most spectra peak in the green with varying amounts of blue and red reflectance, representing non-co-varying concentrations of phytoplankton pigments, CDOM, and suspended inorganic sediments. Several OWTs diverge from this dominant shape; OWT 7 peaks at 510 nm likely due to lower concentrations of chlorophyll *a*, TSS and CDOM that cause R_{rs} to be more dominated by the IOPs of pure water. OWTs 6 and 8 peak at 620 nm, likely due to high concentrations of TSS, which is known to dominate R_{rs} in this region

(Tzortziou et al., 2007), or potentially red tides (Wolny et al., 2020). The shape of the OWTs correspond with the underlying water quality and biogeochemical properties described in Figure 6 and Table S1.

When the OWT classification results were displayed as a map, OWTs were geographically contiguous. In other words, classifying pixels based on their R_{rs} spectra, or color, led to broad clusters of pixels, representing extensive color shifts in the Bay. Seasonal occurrence plots of OWTs demonstrate changes throughout space and time (Figure 5). Throughout space, OWTs are distinctly different in the Upper, Mid, and Lower Bay. In general, the Upper Bay contains OWTs representing turbid waters that have the highest concentrations of water quality parameters (OWTs 6, 8, 12), the Mid Bay contains OWTs that vary in terms of water quality parameters (OWTs 2, 4, 11), and the Lower Bay contains OWTs that represent clear waters with the lowest concentrations of water quality parameters (OWT 1, 5, 7) (Figure 5). However, there is variability to this spatial pattern with some OWTs representing turbid waters occurring in the Lower Bay (OWT 14) and some OWTs representing clear waters occurring in the Upper Bay (OWT 3).

Throughout time, OWTs shift corresponding to seasonal variability of water quality dynamics and freshwater flow (Figures 6 & 7). The OWT that represents the most turbid waters (OWT 6) most commonly occurs in the summer and fall potentially due to sediment resuspension and tidal mixing. The OWT representing the second most turbid waters (OWT 8) is highest in the spring and summer, likely due to mixing from the spring freshet. OWT 6 and 8 encompass the location of the ETM, suggesting that this OWT could potentially be used to track the location and timing of increased particulate concentrations (Sanford et al., 2001). OWT 6 and 8 also commonly occur in the upstream limits of saltwater intrusion of major tributaries, such as the Potomac River, which also contains an ETM (Knebel et al., 1981). OWT 6 and 8 represent waters with the highest nutrient concentrations, suggesting nutrient enrichment can alter OACs and impact OWT location and timing. The OWT representing clear waters with low concentrations of water quality parameters (OWT 3) most commonly occurs in the summer. The occurrence of OWT 3 is similar to OWT 6; however, these OWTs have contrasting underlying optical properties. The occurrence of these OWTs is likely influenced by freshwater flow and nutrient enrichment with OWT 3 occurring during years of low flow, leading to low nutrient concentrations, and OWT 6 occurring during years of high flow, leading to high nutrient concentrations (Figure 6).

The occurrence of OWTs significantly varies with freshwater flow (Figure 7). OWTs representing the most turbid waters have higher occurrence during months of high freshwater flow while OWTs representing clear waters occur more during months of low freshwater flow. This analysis demonstrates how freshwater flow can directly change the most common OWTs in the Bay, signifying how freshwater flow can directly impact underlying water quality properties.

Blended water quality retrievals and future recommendations

The accuracy of seven empirical algorithms designed to estimate concentrations of chlorophyll *a* from measurements of R_{rs} in optically complex waters was analyzed for each OWT. Model F, the normalized difference chlorophyll index (NDCI) algorithm performed best across all OWTs. The selection of wavelengths in this algorithm is sensitive to chlorophyll *a* concentrations in turbid waters and avoids the confounding influence of CDOM and NAP at shorter wavelengths (Mishra & Mishra, 2012). The spectral band difference ratio also helps to eliminate uncertainties associated with atmospheric contributions and seasonal solar azimuth differences (Mishra & Mishra, 2012). NDCI was originally calibrated using *in situ* chlorophyll *a* data collected in several locations including Chesapeake Bay and when applied to satellite data, was within 5-7% of the measured *in situ* values (Mishra & Mishra, 2012). Model F performed best for a range of OWTs representing clear to moderately turbid waters, demonstrating the robustness of this algorithm. The second-best performing chlorophyll *a* algorithm was Model C, a three-band algorithm presented in Dall'Olmo & Gitelson (2005) and Moses et al. (2009). In theory, the combination of three bands alters the model sensitivity by accounting for $R_{rs}(665)$ and $R_{rs}(708)$, which are influenced by NAP, and at $R_{rs}(753)$, which is influenced mainly by particulate backscattering (Moses et al., 2009). Model C performed best for OWTs representing relatively clear waters with low median chlorophyll *a* concentrations, and not as well for more turbid waters. Model D is a two-band NIR-red algorithm and performed best for relatively turbid to highly turbid OWTs. Model E is a three-band

NIR-red algorithm and performed best across a range of OWTs. Finally, Model G, the FLH algorithm, performed best for OWT 13 which represents relatively clear waters with high chlorophyll *a* concentrations. Estimating chlorophyll fluorescence from R_{rs} collected in turbid waters is challenging due to background radiance from scattering particulate matter, which is likely why this algorithm did not perform well often. Model A was never selected as a well performing algorithm since blue-green band-ratio algorithms tend to fail in coastal waters with non co-varying concentrations of IOPs (IOCCG, 2006).

This study only evaluated empirical water quality retrieval algorithms; however, there are other algorithms that would likely produce improved results (IOCCG, 2006). Semi-analytical algorithms use a combination of empiricism and analytical relationships to infer IOPs from R_{rs} and solve the radiative transfer equation to obtain estimates of OACs and tend to perform better in coastal waters (Lee et al., 2002; Zheng & DiGiacomo, 2017; Werdell et al., 2018). Other promising methods include advanced analytical algorithms such as machine learning which involves training a neural net to retrieve simultaneous combinations of water quality retrievals from R_{rs} data (Doerffer & Schiller, 2007; Pahlevan et al., 2020). All water quality retrieval algorithms have uncertainties associated with initial parameterization and atmospheric correction and require high quality IOP measurements to test the consistency between measured and predicted quantities (Zheng & DiGiacomo, 2017; Werdell et al., 2018). While this study analyzed multispectral satellite R_{rs}, future OWT classifications should utilize hyperspectral R_{rs} when available. Future satellite missions such as NASA's Plankton, Atmosphere, Clouds, and ocean Ecosystem (PACE), Surface Biology and Geology (SBG), and Geosynchronous Littoral Imaging and Monitoring Radiometer (GLIMR) will provide unprecedented hyperspectral measurements that will allow for improved characterization of phytoplankton community composition and particle size distributions (Werdell et al., 2019). Hyperspectral R_{rs} will undoubtedly enhance an OWT classification and provide further insight into underlying water quality properties.

Management implications

Satellite remote sensing is the most spatially and temporally comprehensive dataset serving as a useful tool for long-term and extensive water quality monitoring. The average time each mainstem station is sampled using *in situ* and satellite remote sensing methods is demonstrated in Figure 8. Turner et al. (2021) demonstrated how satellite R_{rs} (e.g., color) and band ratios in the Bay have changed over time, contributing to conclusions about phytoplankton bloom phenology, changing suspended particle concentrations, and water clarity. The present study builds upon analyzing water color and employs a classification approach to identify where and when water color is changing. A comprehensive matchup analysis with *in situ* data provides insight into the underlying water quality data associated with individual OWTs. Therefore, the results of this study have the potential to greatly assist Chesapeake Bay water quality monitoring and management.

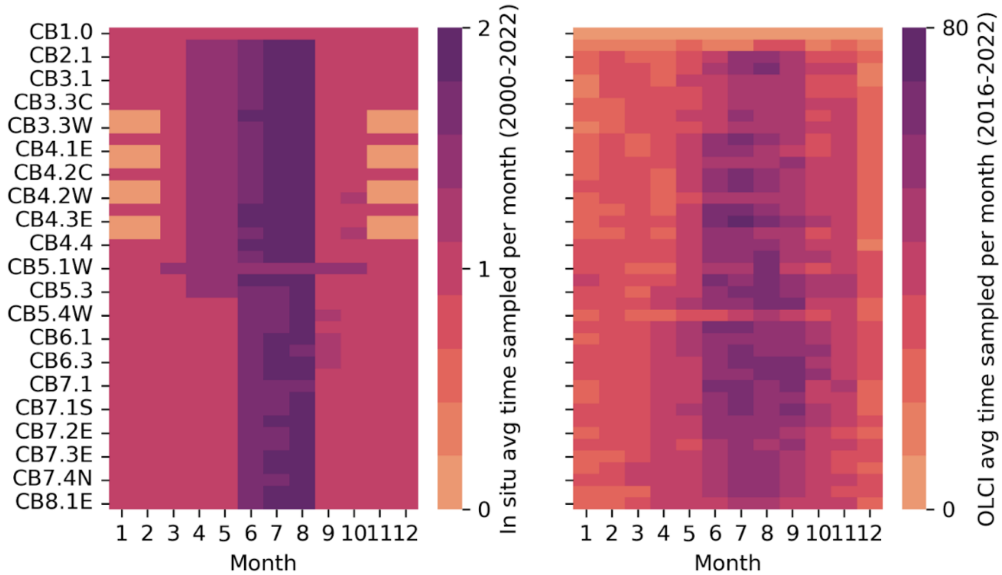


Figure 8. Average number of days sampled at each mainstem station across months between 2000 and 2022 using *in situ* and satellite remote sensing methods.

The OWT classification delineates the optical diversity of the Chesapeake Bay through space and time, with many OWTs occurring in nearshore waters and tributaries. The CBP tidal water quality monitoring program whose stations are predomantly in the main stem may therefore not be representative of this optical diversity and as such may be missing ecologically and biogeochemically areas that could be of management concern. The timing and location of OWTs can be used to inform more focused sampling, particularly in areas containing high optical diversity. Additionally, an OWT classification could assist with assessing water quality standards (Tango & Batiuk, 2013). For example, OWTs 6, 8, and 14 represent turbid waters with the highest median K_d (medians = 2.27, 1.92, 1.52, respectively) and could represent areas that are not meeting a numerical water clarity criteria (Tango & Batiuk, 2013). The high spatial and temporal resolution of an OWT classification could assist with developing TMDLs and assessing progress in water quality response. Furthermore, an OWT classification could potentially assist with algal bloom monitoring. The location and timing of a known harmful algal bloom (*Margalefidinium polykrikoides*) in the York River (Wolny et al., 2020) was compared to the OWT results. The HAB is represented by OWT 15, which represents waters with the highest median chlorophyll *a* concentration. This finding demonstrates that an OWT classification has the potential to identify near real-time locations containing high chlorophyll which could be locations of potential HABs.

Satellite R_{rs} also has the potential to be incorporated into numerical models. The CBP has a long history of using numerical models as a planning tool to inform management decisions toward Bay restoration (Hood et al., 2021). Specifically, the CBP modeling system is used to 1) set nutrient and sediment reduction targets, 2) estimate nutrient and sediment reduction plans to meet the targets (e.g., total maximum daily loads, TMDLs), and 3) quantify progress of Bay-wide restoration goals (Hood et al., 2021). Since the CBP modeling system greatly influences Bay policy and funding decisions, it is necessary to develop and evaluate model outputs with known data to be scientifically defensible. *In situ* observations are typically used for model development and evaluation (Xu & Hood, 2006); however, they can be sparse in time and space due to the prohibitive expense of collecting them. Studies have shown how satellite data can be assimilated into ecosystem models to reduce error and recover optimal parameter sets (Friedrichs, 2002; Jones et al., 2016; Friedrichs & Kaufman, 2019; IOCCG, 2020). Jones et al. (2016) demonstrated how assimilating R_{rs} instead of derived water quality products that may have compounded uncertainties can significantly improve model predictive skill for optically complex waters. Assimilating near-real time synoptic satellite R_{rs} data into numerical models is an active area of research (Gregg et al., 2009; Ciavatta et al., 2014; IOCCG, 2020) and has potential to significantly enhance Chesapeake Bay water quality management.

Conclusion

This study was the first of its kind to apply an optical water type (OWT) classification based on normalized remote sensing reflectance to study spatial and temporal dynamics of ecologically and biogeochemically important constituents in the Chesapeake Bay. Fifteen OWTs were statistically identified from a *k*-means clustering algorithm ranging in shape due to co-varying optically active constituents. A comprehensive matchup analysis with *in situ* water quality data characterized each OWT in terms of underlying water quality data. The location and timing of OWTs with high concentrations of optically active constituents can be of particular interest for water quality management and help inform more focused sampling. The accuracy of seven empirical algorithms designed to estimate concentrations of chlorophyll *a* from measurements of R_{rs} was analyzed for each OWT. The best performing algorithms can be applied in a dynamic switching algorithm to produce an improved blended chlorophyll *a* product for the Chesapeake Bay.

Supplementary material

Table S1. Summary statistics of *in situ* water quality data corresponding to each OWT.

		Chl *a* (µg/L)	**TSS (mg/L)**	**K_d (m-1)**	**Z_{SD} (m)**	**TN (mg/L)**	**TP (mg/L)**	**Inorganic particles (mg/L)**	**Organic particles (mg/L)**	**Chl:C (ug/mg)**	**PN:PP (mol:mol)**
OWT 1 (n=39)	count	39	33	15	39	28	31	19	19	19	28
	0.25	3.99	5.48	0.63	0.85	0.44	0.02	2.99	2.41	2.71	18.62
	median	5.62	7.00	0.73	1.10	0.59	0.03	5.00	2.80	4.55	23.07
	0.75	8.97	9.00	0.90	1.30	0.91	0.04	6.90	3.10	5.49	25.83
	mean	6.84	7.43	0.88	1.08	0.66	0.04	4.93	2.94	4.44	22.61
	std	3.79	2.80	0.49	0.31	0.34	0.04	2.42	0.84	1.97	6.36
	min, max	1.98, 17.9	2.40, 14.8	0.49, 2.46	0.20, 1.60	0.23, 1.41	0.01, 0.24	1.58, 9.20	2.00, 4.8	2.05, 9.34	10.1, 41.43
OWT 2 (n=32)	count	30	23	8	32	14	23	17	17	17	14
	0.25	3.85	6.90	1.18	0.80	0.51	0.03	4.00	2.50	2.75	14.73
	median	5.77	8.50	1.48	0.90	0.86	0.04	5.40	3.00	4.67	16.70
	0.75	9.45	13.43	1.78	1.03	1.01	0.05	8.50	4.00	7.88	22.75
	mean	6.73	11.76	1.54	0.96	0.82	0.04	8.62	3.71	4.81	18.70
	std	4.30	9.07	0.47	0.27	0.36	0.02	8.48	2.14	3.14	7.14
	min, max	0.43, 17.73	5.00, 43.00	0.95, 2.36	0.50, 1.70	0.38, 1.52	0.02, 0.08	2.00, 35.0	1.00, 8.00	0.60, 9.53	6.97, 35.3
OWT 3 (n=127)	count	117	107	105	127	114	98	44	44	44	118
	0.25	4.39	4.56	0.48	1.20	0.34	0.02	2.00	2.43	2.94	21.07
	median	6.65	5.20	0.68	1.50	0.44	0.02	2.61	2.88	3.81	24.12
	0.75	8.33	6.39	0.88	1.90	0.54	0.03	3.84	3.61	4.88	27.22
	mean	7.46	5.78	0.74	1.60	0.45	0.03	3.54	3.09	4.45	24.45
	std	5.40	3.03	0.35	0.53	0.15	0.01	3.57	1.01	3.37	5.14
	min, max	1.66, 49.0	2.40, 32.0	0.20, 2.11	0.60, 3.00	0.14, 0.91	0.01, 0.10	0.97, 24.4	1.80, 7.66	1.84, 24.5	11.2, 43.3
OWT 4 (n=387)	count	369	333	299	386	339	282	177	179	179	354
	0.25	5.23	5.10	0.51	1.20	0.34	0.02	2.40	2.60	3.10	19.40
	median	6.88	6.00	0.67	1.40	0.39	0.03	3.68	3.10	4.22	22.13
	0.75	9.18	7.20	0.91	1.60	0.50	0.03	4.72	3.50	5.36	26.18
	mean	7.93	6.34	0.77	1.45	0.45	0.03	3.89	3.13	4.63	23.15
	std	7.28	1.96	0.37	0.40	0.19	0.01	1.91	0.85	2.09	5.47

	min, max	1.50, 116	2.6, 17.1	0.14, 2.35	0.60, 3.50	0.16, 1.67	0.01, 0.08	0.40, 11.8	0.85, 6.00	1.53, 13.9	12.9, 45.3
OWT 5 (n=115)	count	104	96	60	115	83	82	52	51	51	83
	0.25	5.66	5.00	0.64	1.10	0.39	0.02	2.40	2.46	3.28	19.72
	median	7.37	6.00	0.74	1.30	0.48	0.03	3.45	3.10	4.12	24.45
	0.75	10.06	7.43	0.87	1.60	0.63	0.04	5.44	4.00	5.04	29.12
	mean	8.31	8.16	0.75	1.33	0.57	0.03	7.09	3.58	4.38	25.55
	std	4.31	11.26	0.24	0.34	0.31	0.02	13.57	2.59	1.90	8.09
	min, max	1.97, 21.2	3.00, 102	0.22, 1.32	0.50, 2.20	0.25, 2.19	0.01, 0.10	1.00, 82.0	1.00, 20.0	0.34, 9.30	9.89, 57.7
OWT 6 (n=32)	count	31	29	7	32	26	29	21	21	20	26
	0.25	3.94	10.80	2.20	0.40	0.85	0.04	7.60	3.60	1.70	10.00
	median	7.48	16.00	2.27	0.50	1.01	0.05	14.30	4.65	2.70	15.44
	0.75	9.79	21.00	2.69	0.63	1.48	0.07	21.20	6.70	5.33	19.29
	mean	8.75	18.61	2.40	0.59	1.14	0.05	16.03	5.17	3.40	15.42
	std	7.44	11.67	0.33	0.33	0.45	0.02	11.17	1.95	2.41	6.04
	min, max	0.85, 32.8	2.80, 55.5	1.94, 2.78	0.20, 1.60	0.51, 1.96	0.02 ,0.11	3.00, 47.0	2.72, 8.50	0.57, 9.37	7.22, 27.5
OWT 7 (n=23)	count	21	22	21	22	23	21	7	7	7	23
	0.25	4.27	3.87	0.56	1.05	0.45	0.02	1.90	2.76	2.90	22.84
	median	7.74	5.59	0.77	1.45	0.55	0.03	3.46	2.90	3.44	28.97
	0.75	10.01	7.82	0.98	1.93	0.61	0.03	4.59	3.21	4.83	36.15
	mean	9.25	6.13	0.80	1.58	0.56	0.03	3.39	2.94	4.35	29.44
	std	6.31	2.70	0.33	0.61	0.20	0.02	1.72	0.57	2.54	7.44
	min, max	3.04, 23.92	2.80, 12.0	0.39, 1.62	0.8, 2.70	0.26, 1.16	0.01, 0.08	1.42, 5.90	2.00, 3.75	2.1, 9.46	16.7, 44.0
OWT 8 (n=71)	count	65	63	28	71	51	55	43	43	43	54
	0.25	4.91	11.10	1.42	0.40	0.72	0.04	6.73	3.25	2.44	10.57
	median	8.54	14.40	1.92	0.60	1.03	0.05	10.00	5.00	3.94	16.19
	0.75	15.50	20.05	2.33	0.70	1.36	0.06	18.40	7.00	5.33	24.27
	mean	14.98	19.15	1.90	0.60	1.06	0.05	15.41	6.06	4.28	17.94
	std	20.19	16.73	0.58	0.28	0.39	0.02	16.13	4.41	2.46	8.56
	min, max	1.42, 107	5.60, 94.0	0.77, 3.09	0.20, 2.00	0.46, 2.04	0.02, 0.09	0.40, 80.0	2.00, 26.0	0.65, 13.4	6.8, 39.4
OWT 9 (n=69)	count	66	63	23	69	57	54	37	37	37	56
	0.25	7.34	4.95	0.75	0.90	0.49	0.02	3.00	2.80	3.73	17.12
	median	8.89	7.10	0.91	1.10	0.58	0.03	6.40	3.00	5.46	21.90
	0.75	11.32	9.89	1.03	1.40	0.79	0.05	8.40	4.35	7.08	29.85
	mean	9.59	9.03	0.96	1.17	0.66	0.04	7.50	3.48	5.74	23.47
	std	4.00	6.96	0.37	0.37	0.29	0.01	7.34	1.31	2.37	8.43
	min, max	3.54, 25.2	2.86, 47.6	0.40, 2.03	0.30, 2.00	0.21, 1.65	0.02, 0.08	1.00, 41.2	1.00, 6.40	1.84, 12.6	8.5, 43.3
OWT 10	count	166	150	125	176	152	137	84	84	82	157
	0.25	6.41	5.51	0.63	1.00	0.38	0.02	3.08	2.52	3.75	19.72
	median	8.94	7.13	0.79	1.25	0.46	0.03	4.71	3.33	5.23	22.80
	0.75	11.87	8.79	0.98	1.50	0.60	0.04	6.56	4.00	6.63	27.94

	mean	10.21	7.62	0.83	1.29	0.50	0.03	5.20	3.41	5.18	23.65
	std	6.14	3.13	0.32	0.41	0.19	0.01	3.08	1.12	1.85	6.57
	min, max	1.92, 48.9	2.98, 23.0	0.13, 1.83	0.60, 3.20	0.19, 1.29	0.01, 0.08	0.90, 18.3	1.60, 6.23	1.75, 10.3	7.78, 41.1
OWT 11 (n=53)	count	51	47	16	53	43	39	31	31	31	42
	0.25	6.28	8.60	1.21	0.50	0.57	0.03	6.25	3.10	3.39	15.24
	median	9.19	10.30	1.42	0.70	0.78	0.04	8.93	4.40	4.56	20.01
	0.75	13.95	14.80	1.47	1.00	1.10	0.05	14.85	5.00	6.57	23.74
	mean	11.24	13.02	1.37	0.77	0.87	0.04	11.11	4.61	5.11	19.99
	std	7.49	9.18	0.32	0.34	0.40	0.01	8.38	2.24	2.46	6.57
	min, max	1.07, 41.4	3.50, 58.0	0.75, 2.09	0.30, 2.00	0.31, 1.90	0.02, 0.07	0.80, 44.0	2.00, 14.0	1.20, 10.34	9.99, 35.1
OWT 12 (n=47)	count	44	39	9	47	32	33	31	31	31	33
	0.25	6.56	6.73	1.29	0.70	0.50	0.03	4.00	3.63	3.19	14.47
	median	9.77	9.60	1.41	0.90	0.66	0.04	5.50	4.40	4.43	20.31
	0.75	16.30	13.60	1.66	1.00	0.81	0.05	9.05	5.65	6.80	26.14
	mean	10.94	11.79	1.48	0.87	0.71	0.04	8.47	4.69	4.91	20.81
	std	6.22	8.63	0.45	0.26	0.29	0.02	8.71	1.72	2.12	7.15
	min, max	1.07, 29.9	4.20, 53.0	0.71, 2.38	0.50, 1.60	0.319, 1.61	0.02, 0.09	0.40, 47.0	2.80, 11.0	1.71, 9.50	9.55, 37.1
OWT 13 (n=144)	count	133	129	107	144	120	108	58	58	57	129
	0.25	7.93	4.80	0.80	0.90	0.53	0.03	2.29	2.73	5.11	20.80
	median	10.70	6.90	0.98	1.10	0.63	0.03	3.89	3.20	6.69	23.71
	0.75	14.28	9.10	1.15	1.50	0.75	0.04	6.26	4.00	7.74	26.81
	mean	11.87	7.58	1.03	1.19	0.66	0.03	5.33	3.59	6.82	24.12
	std	5.99	3.83	0.41	0.39	0.21	0.01	4.37	1.32	2.77	5.57
	min, max	2.14, 40.9	2.50, 23.0	0.36, 3.25	0.50, 2.20	0.33, 1.35	0.01, 0.08	0.80, 18.2	1.10, 7.20	2.83, 21.0	5.75, 47.2
OWT 14 (n=86)	count	82	76	40	86	77	74	33	33	33	79
	0.25	8.68	7.52	1.32	0.40	0.68	0.04	8.00	4.40	2.81	16.08
	median	14.95	11.20	1.52	0.70	0.86	0.05	15.20	6.00	4.72	21.44
	0.75	22.99	20.10	1.89	0.80	1.07	0.06	19.30	8.00	6.92	25.07
	mean	24.22	16.57	1.68	0.69	1.05	0.05	17.59	7.90	5.39	21.51
	std	47.15	14.48	0.68	0.33	0.74	0.02	13.88	7.39	3.78	7.35
	min, max	1.60, 373	3.60, 70.0	0.03, 3.36	0.20, 1.90	0.28, 5.80	0.02, 0.12	1.76, 60.7	2.00, 37.0	1.39, 20.2	7.27, 44.9
OWT 15 (n=110)	count	105	103	65	110	91	88	40	40	40	92
	0.25	10.41	6.25	1.03	0.70	0.64	0.03	2.60	3.38	4.79	20.04
	median	16.29	8.00	1.29	0.90	0.71	0.04	4.20	4.00	7.14	24.50
	0.75	23.07	10.40	1.51	1.10	0.85	0.06	6.55	5.00	9.84	26.84
	mean	18.68	8.77	1.27	0.94	0.76	0.05	4.99	4.16	7.25	24.31
	std	11.70	3.65	0.36	0.35	0.21	0.02	3.33	1.23	3.20	6.15
	min, max	0.64, 66.0	2.8, 20.4	0.45, 2.35	0.40, 2.80	0.40, 1.54	0.02, 0.12	1.00, 16.9	2.00, 8.00	0.46, 15.1	11.4, 50.5

Table S2. *p*-value results from a pairwise Tukey's honest significance test for TN, TP, Chl *a*, TSS, and K_d. Shaded values are *p*-values > 0.05, representing OWTs that contain statistically similar means.

TN

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1		0.97	0.06	0.02	0.98	0.00	1.00	0.00	1.00	0.31	0.21	1.00	1.00	0.00	0.97
2	0.97		0.00	0.00	0.22	0.09	0.41	0.32	0.93	0.01	1.00	1.00	0.90	0.32	1.00
3	0.06	0.00		1.00	0.30	0.00	0.97	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00
4	0.02	0.00	1.00		0.06	0.00	0.92	0.00	0.00	0.89	0.00	0.00	0.00	0.00	0.00
5	0.98	0.22	0.30	0.06		0.00	1.00	0.00	0.89	0.93	0.00	0.61	0.65	0.00	0.00
6	0.00	0.09	0.00	0.00	0.00		0.00	1.00	0.00	0.00	0.03	0.00	0.00	1.00	0.00
7	1.00	0.41	0.97	0.92	1.00	0.00		0.00	0.99	1.00	0.00	0.87	0.97	0.00	0.18
8	0.00	0.32	0.00	0.00	0.00	1.00	0.00		0.00	0.00	0.16	0.00	0.00	1.00	0.00
9	1.00	0.93	0.00	0.00	0.89	0.00	0.99	0.00		0.03	0.04	1.00	1.00	0.00	0.80
10	0.31	0.01	1.00	0.89	0.93	0.00	1.00	0.00	0.03		0.00	0.02	0.00	0.00	0.00
11	0.21	1.00	0.00	0.00	0.00	0.03	0.00	0.16	0.04	0.00		0.56	0.01	0.13	0.81
12	1.00	1.00	0.00	0.00	0.61	0.00	0.87	0.00	1.00	0.02	0.56		1.00	0.00	1.00
13	1.00	0.90	0.00	0.00	0.65	0.00	0.97	0.00	1.00	0.00	0.01	1.00		0.00	0.51
14	0.00	0.32	0.00	0.00	0.00	1.00	0.00	1.00	0.00	0.00	0.13	0.00	0.00		0.00
15	0.97	1.00	0.00	0.00	0.00	0.00	0.18	0.00	0.80	0.00	0.81	1.00	0.51	0.00	

TP

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1		1.00	0.01	0.02	0.92	0.10	0.85	0.22	1.00	0.42	1.00	1.00	1.00	0.00	0.40
2	1.00		0.00	0.00	0.37	0.77	0.37	0.96	0.99	0.06	1.00	1.00	0.80	0.30	1.00
3	0.01	0.00		1.00	0.15	0.00	1.00	0.00	0.00	0.54	0.00	0.00	0.00	0.00	0.00
4	0.02	0.00	1.00		0.29	0.00	1.00	0.00	0.00	0.82	0.00	0.00	0.00	0.00	0.00
5	0.92	0.37	0.15	0.29		0.00	1.00	0.00	0.98	1.00	0.32	0.30	1.00	0.00	0.00
6	0.10	0.77	0.00	0.00	0.00		0.00	1.00	0.01	0.00	0.30	0.47	0.00	1.00	0.99
7	0.85	0.37	1.00	1.00	1.00	0.00		0.00	0.93	1.00	0.39	0.35	0.99	0.00	0.00
8	0.22	0.96	0.00	0.00	0.00	1.00	0.00		0.01	0.00	0.58	0.79	0.00	0.99	1.00
9	1.00	0.99	0.00	0.00	0.98	0.01	0.93	0.01		0.44	1.00	0.99	1.00	0.00	0.02
10	0.42	0.06	0.54	0.82	1.00	0.00	1.00	0.00	0.44		0.02	0.03	0.60	0.00	0.00
11	1.00	1.00	0.00	0.00	0.32	0.30	0.39	0.58	1.00	0.02		1.00	0.82	0.02	0.82
12	1.00	1.00	0.00	0.00	0.30	0.47	0.35	0.79	0.99	0.03	1.00		0.78	0.06	0.94
13	1.00	0.80	0.00	0.00	1.00	0.00	0.99	0.00	1.00	0.60	0.82	0.78		0.00	0.00
14	0.00	0.30	0.00	0.00	0.00	1.00	0.00	0.99	0.00	0.00	0.02	0.06	0.00		0.61
15	0.40	1.00	0.00	0.00	0.00	0.99	0.00	1.00	0.02	0.00	0.82	0.94	0.00	0.61	

Chl a

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1		1.00	1.00	1.00	1.00	1.00	1.00	0.18	1.00	0.99	0.98	0.99	0.79	0.00	0.00
2	1.00		1.00	1.00	1.00	1.00	1.00	0.29	1.00	1.00	0.99	0.99	0.88	0.00	0.00
3	1.00	1.00		1.00	1.00	1.00	1.00	0.03	1.00	0.94	0.95	0.98	0.41	0.00	0.00
4	1.00	1.00	1.00		1.00	1.00	1.00	0.01	1.00	0.91	0.96	0.99	0.23	0.00	0.00
5	1.00	1.00	1.00	1.00		1.00	1.00	0.13	1.00	1.00	1.00	1.00	0.81	0.00	0.00
6	1.00	1.00	1.00	1.00	1.00		1.00	0.75	1.00	1.00	1.00	1.00	1.00	0.00	0.03
7	1.00	1.00	1.00	1.00	1.00	1.00		0.94	1.00	1.00	1.00	1.00	1.00	0.00	0.21
8	0.18	0.29	0.03	0.01	0.13	0.75	0.94		0.63	0.53	0.98	0.98	0.98	0.01	0.93
9	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.63		1.00	1.00	1.00	1.00	0.00	0.00
10	0.99	1.00	0.94	0.91	1.00	1.00	1.00	0.53	1.00		1.00	1.00	1.00	0.00	0.00
11	0.98	0.99	0.95	0.96	1.00	1.00	1.00	0.98	1.00	1.00		1.00	1.00	0.00	0.10
12	0.99	0.99	0.98	0.99	1.00	1.00	1.00	0.98	1.00	1.00	1.00		1.00	0.00	0.11
13	0.79	0.88	0.41	0.23	0.81	1.00	1.00	0.98	1.00	1.00	1.00	1.00		0.00	0.01
14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00		0.28
15	0.00	0.00	0.00	0.00	0.00	0.03	0.21	0.93	0.00	0.00	0.10	0.11	0.01	0.28	

TSS

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1		0.66	1.00	1.00	1.00	0.00	1.00	0.00	1.00	1.00	0.05	0.41	1.00	0.00	1.00
2	0.66		0.03	0.04	0.70	0.05	0.36	0.00	0.97	0.40	1.00	1.00	0.40	0.25	0.90
3	1.00	0.03		1.00	0.55	0.00	1.00	0.00	0.23	0.79	0.00	0.00	0.85	0.00	0.16
4	1.00	0.04	1.00		0.68	0.00	1.00	0.00	0.31	0.90	0.00	0.00	0.95	0.00	0.17
5	1.00	0.70	0.55	0.68		0.00	1.00	0.00	1.00	1.00	0.01	0.35	1.00	0.00	1.00
6	0.00	0.05	0.00	0.00	0.00		0.00	1.00	0.00	0.00	0.07	0.01	0.00	0.99	0.00
7	1.00	0.36	1.00	1.00	1.00	0.00		0.00	0.96	1.00	0.02	0.18	1.00	0.00	0.97
8	0.00	0.00	0.00	0.00	0.00	1.00	0.00		0.00	0.00	0.00	0.00	0.00	0.74	0.00
9	1.00	0.97	0.23	0.31	1.00	0.00	0.96	0.00		0.99	0.21	0.87	0.99	0.00	1.00
10	1.00	0.40	0.79	0.90	1.00	0.00	1.00	0.00	0.99		0.00	0.09	1.00	0.00	1.00
11	0.05	1.00	0.00	0.00	0.01	0.07	0.02	0.00	0.21	0.00		1.00	0.00	0.34	0.06
12	0.41	1.00	0.00	0.00	0.35	0.01	0.18	0.00	0.87	0.09	1.00		0.09	0.06	0.65
13	1.00	0.40	0.85	0.95	1.00	0.00	1.00	0.00	0.99	1.00	0.00	0.09		0.00	1.00
14	0.00	0.25	0.00	0.00	0.00	0.99	0.00	0.74	0.00	0.00	0.34	0.06	0.00		0.00
15	1.00	0.90	0.16	0.17	1.00	0.00	0.97	0.00	1.00	1.00	0.06	0.65	1.00	0.00	

Kd

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1		0.01	0.99	1.00	1.00	0.00	1.00	0.00	1.00	1.00	0.04	0.02	0.99	0.00	0.04
2	0.01		0.00	0.00	0.00	0.00	0.00	0.61	0.02	0.00	1.00	1.00	0.03	1.00	0.87
3	0.99	0.00		1.00	1.00	0.00	1.00	0.00	0.44	0.84	0.00	0.00	0.00	0.00	0.00
4	1.00	0.00	1.00		1.00	0.00	1.00	0.00	0.62	0.96	0.00	0.00	0.00	0.00	0.00
5	1.00	0.00	1.00	1.00		0.00	1.00	0.00	0.68	0.99	0.00	0.00	0.00	0.00	0.00
6	0.00	0.00	0.00	0.00	0.00		0.00	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00
7	1.00	0.00	1.00	1.00	1.00	0.00		0.00	0.99	1.00	0.00	0.00	0.45	0.00	0.00
8	0.00	0.61	0.00	0.00	0.00	0.14	0.00		0.00	0.00	0.00	0.23	0.00	0.61	0.00
9	1.00	0.02	0.44	0.62	0.68	0.00	0.99	0.00		0.99	0.08	0.05	1.00	0.00	0.06
10	1.00	0.00	0.84	0.96	0.99	0.00	1.00	0.00	0.99		0.00	0.00	0.01	0.00	0.00
11	0.04	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.08	0.00		1.00	0.08	0.29	1.00
12	0.02	1.00	0.00	0.00	0.00	0.00	0.00	0.23	0.05	0.00	1.00		0.07	0.99	0.98
13	0.99	0.03	0.00	0.00	0.00	0.00	0.45	0.00	1.00	0.01	0.08	0.07		0.00	0.01
14	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.61	0.00	0.00	0.29	0.99	0.00		0.00
15	0.04	0.87	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.00	1.00	0.98	0.01	0.00	

Table S3. Performance metrics comparing modeled and *in situ* chlorophyll *a* for each OWT.

OWT #	Model	n	slope	int	r	mae	bias	rmse	r^2	*a*	*b*	*c*
1	1	39	-7.34	15.11	-0.03	8.68	0.22	8.86	0.00			
1	2	39	0.06	0.72	0.06	0.19	0.00	0.24	0.00	-0.40	1.04	
1	3	39	0.21	0.61	0.21	0.19	0.00	0.23	0.04	0.00	0.65	
1	4	39	0.20	0.62	0.20	0.19	0.00	0.23	0.04	29.33	-39.35	13.93
1	5	39	0.25	0.58	0.25	0.19	0.00	0.23	0.06	698.68	128.33	6.62
1	6	39	0.19	0.62	0.19	0.19	0.00	0.23	0.04	2.80	20.88	52.72
1	7	39	0.23	0.60	0.23	0.19	0.00	0.23	0.05	1.00	-959.07	
2	1	30	38.87	-6.74	0.33	20.58	0.69	24.45	0.11			
2	2	30	0.13	0.63	0.13	0.26	0.00	0.34	0.02	-0.75	1.21	
2	3	30	0.38	0.45	0.38	0.26	0.00	0.32	0.15	0.00	0.35	
2	4	30	0.16	0.61	0.16	0.26	0.00	0.34	0.03	9.53	-13.17	5.22
2	5	30	0.13	0.63	0.13	0.26	0.00	0.34	0.02	409.95	70.33	3.70
2	6	30	0.14	0.62	0.14	0.26	0.00	0.34	0.02	0.94	3.14	9.72
2	7	30	0.08	0.66	0.08	0.26	0.00	0.34	0.01	0.83	-270.63	
3	1	117	-1.43	5.04	-0.04	3.08	0.03	3.11	0.00			
3	2	117	0.35	0.53	0.35	0.16	0.00	0.21	0.12	2.84	-0.93	
3	3	117	0.45	0.44	0.45	0.16	0.00	0.20	0.20	0.00	0.62	
3	4	117	0.44	0.45	0.44	0.16	0.00	0.20	0.20	-38.63	51.40	-16.14
3	5	117	0.22	0.63	0.22	0.17	0.00	0.22	0.05	505.86	109.54	6.69
3	6	117	0.44	0.45	0.44	0.16	0.00	0.20	0.20	-1.68	-26.10	-64.85
3	7	117	0.32	0.55	0.32	0.17	0.00	0.21	0.10	1.05	-1864.10	
4	1	369	1.19	3.62	0.09	3.78	0.01	3.79	0.01			
4	2	369	0.35	0.55	0.35	0.14	0.00	0.20	0.12	2.63	-0.83	
4	3	369	0.23	0.64	0.23	0.15	0.00	0.20	0.06	0.00	0.70	
4	4	369	0.37	0.53	0.37	0.14	0.00	0.19	0.13	-25.18	34.73	-11.04
4	5	369	0.24	0.64	0.24	0.15	0.00	0.20	0.06	165.70	42.06	3.43
4	6	369	0.37	0.53	0.37	0.14	0.00	0.19	0.14	-0.48	-15.61	-43.01
4	7	369	0.27	0.62	0.27	0.15	0.00	0.20	0.07	1.07	-1474.68	
5	1	104	-2.33	7.89	-0.06	5.02	0.05	5.05	0.00			
5	2	104	0.45	0.47	0.45	0.17	0.00	0.20	0.20	2.67	-0.92	
5	3	104	0.22	0.67	0.22	0.18	0.00	0.22	0.05	0.00	0.74	
5	4	104	0.45	0.47	0.45	0.17	0.00	0.20	0.20	6.67	-6.30	2.09
5	5	104	0.40	0.52	0.40	0.17	0.00	0.21	0.16	134.01	35.40	3.03
5	6	104	0.45	0.47	0.45	0.17	0.00	0.20	0.21	2.34	11.16	18.81
5	7	104	0.22	0.67	0.22	0.18	0.00	0.22	0.05	1.06	-1138.02	
6	1	31	36.83	-0.60	0.09	27.90	0.90	31.18	0.01			
6	2	31	0.25	0.60	0.25	0.28	0.00	0.37	0.06	0.78	0.01	
6	3	31	0.29	0.57	0.29	0.28	0.00	0.36	0.08	0.00	0.66	
6	4	31	0.48	0.42	0.48	0.26	0.00	0.33	0.23	6.01	-12.16	6.86
6	5	31	0.41	0.47	0.41	0.27	0.00	0.34	0.17	17.74	-0.31	0.74
6	6	31	0.51	0.39	0.51	0.26	0.00	0.33	0.26	0.69	0.20	32.17
6	7	31	0.07	0.74	0.07	0.30	0.00	0.38	0.00	0.78	-70.28	

7	1	21	-1.47	3.77	-0.07	1.60	0.08	1.67	0.00			
7	2	21	0.37	0.56	0.37	0.20	0.00	0.25	0.13	4.48	-1.66	
7	3	21	0.62	0.34	0.62	0.17	0.00	0.21	0.38	0.00	0.66	
7	4	21	0.37	0.56	0.37	0.20	0.00	0.25	0.13	-9.61	15.48	-4.80
7	5	21	0.41	0.52	0.41	0.21	0.00	0.25	0.17	764.36	186.15	12.12
7	6	21	0.37	0.56	0.37	0.20	0.00	0.25	0.13	2.15	3.66	-3.45
7	7	21	0.15	0.75	0.15	0.23	0.00	0.27	0.02	0.94	-589.28	
8	1	65	-16.51	33.49	-0.07	16.94	0.26	18.39	0.00			
8	2	65	0.35	0.62	0.35	0.30	0.00	0.40	0.12	1.05	-0.06	
8	3	65	0.42	0.55	0.42	0.31	0.00	0.39	0.17	0.00	0.60	
8	4	65	0.48	0.50	0.48	0.28	0.00	0.38	0.23	6.20	-11.18	5.84
8	5	65	0.44	0.53	0.44	0.29	0.00	0.39	0.20	33.04	2.22	0.87
8	6	65	0.48	0.49	0.48	0.28	0.00	0.38	0.23	0.87	2.99	25.18
8	7	65	0.18	0.77	0.18	0.32	0.00	0.42	0.03	0.93	-183.30	
9	1	66	3.10	3.59	0.04	5.58	0.08	5.61	0.00			
9	2	66	0.32	0.64	0.32	0.13	0.00	0.17	0.10	1.40	-0.10	
9	3	66	0.13	0.82	0.13	0.13	0.00	0.17	0.02	0.00	0.90	
9	4	66	0.32	0.64	0.32	0.13	0.00	0.16	0.11	-3.55	6.71	-2.08
9	5	66	0.34	0.62	0.34	0.13	0.00	0.16	0.12	-12.79	3.76	1.31
9	6	66	0.32	0.64	0.32	0.13	0.00	0.17	0.10	1.18	1.16	-3.35
9	7	66	0.24	0.72	0.24	0.13	0.00	0.17	0.06	1.12	-908.57	
10	1	166	1.76	3.10	0.12	3.82	0.02	3.85	0.01			
10	2	166	0.19	0.77	0.19	0.17	0.00	0.22	0.03	1.13	0.13	
10	3	166	0.34	0.62	0.34	0.16	0.00	0.21	0.12	0.00	0.84	
10	4	166	0.19	0.77	0.19	0.17	0.00	0.22	0.04	4.47	-5.34	2.47
10	5	166	0.05	0.90	0.05	0.17	0.00	0.22	0.00	-19.45	-2.17	0.91
10	6	166	0.19	0.77	0.19	0.17	0.00	0.22	0.03	1.43	4.33	8.18
10	7	166	0.33	0.63	0.33	0.16	0.00	0.21	0.11	1.25	-1682.81	
11	1	51	-11.74	24.31	-0.18	12.06	0.24	12.54	0.03			
11	2	51	0.33	0.65	0.33	0.22	0.00	0.27	0.11	0.91	0.16	
11	3	51	0.26	0.72	0.26	0.22	0.00	0.28	0.07	0.00	0.81	
11	4	51	0.34	0.63	0.34	0.22	0.00	0.27	0.12	-2.22	5.01	-1.70
11	5	51	0.33	0.65	0.33	0.22	0.00	0.27	0.11	-22.43	2.05	1.09
11	6	51	0.35	0.63	0.35	0.22	0.00	0.27	0.12	1.08	1.05	-6.26
11	7	51	0.23	0.74	0.23	0.23	0.00	0.28	0.05	0.97	-216.37	
12	1	44	-3.41	12.22	-0.01	8.00	0.18	8.07	0.00			
12	2	44	0.40	0.57	0.40	0.22	0.00	0.27	0.16	1.73	-0.43	
12	3	44	0.32	0.65	0.32	0.22	0.00	0.28	0.10	0.00	0.72	
12	4	44	0.42	0.55	0.42	0.21	0.00	0.27	0.18	-8.36	15.16	-5.79
12	5	44	0.41	0.57	0.41	0.22	0.00	0.27	0.17	-132.33	-9.03	0.94
12	6	44	0.42	0.55	0.42	0.21	0.00	0.27	0.18	1.07	-1.56	-19.49
12	7	44	0.39	0.58	0.39	0.22	0.00	0.27	0.16	1.16	-1039.52	
13	1	133	-2.19	7.59	-0.03	4.32	0.03	4.34	0.00			
13	2	133	0.32	0.70	0.32	0.15	0.00	0.20	0.10	1.73	-0.36	
13	3	133	0.16	0.86	0.16	0.16	0.00	0.20	0.03	0.00	0.97	
13	4	133	0.32	0.70	0.32	0.15	0.00	0.20	0.10	-1.27	3.78	-1.19

13	5	133	0.21	0.81	0.21	0.15	0.00	0.20	0.04	36.34	8.18	1.39
13	6	133	0.32	0.70	0.32	0.15	0.00	0.20	0.10	1.33	2.82	0.00
13	7	133	0.33	0.69	0.33	0.15	0.00	0.19	0.11	1.46	-2749.42	
14	1	82	-2.80	11.31	-0.21	6.93	0.08	7.04	0.05			
14	2	82	0.36	0.74	0.36	0.27	0.00	0.36	0.13	1.13	-0.05	
14	3	82	0.22	0.89	0.22	0.27	0.00	0.37	0.05	0.00	1.00	
14	4	82	0.54	0.53	0.54	0.25	0.00	0.32	0.29	7.81	-16.43	9.68
14	5	82	0.55	0.52	0.55	0.25	0.00	0.32	0.30	19.00	-0.62	1.09
14	6	82	0.53	0.54	0.53	0.26	0.00	0.33	0.28	1.04	-1.81	43.53
14	7	82	0.09	1.05	0.09	0.28	0.00	0.38	0.01	1.13	-70.04	
15	1	105	-1.73	8.53	-0.07	5.30	0.05	5.33	0.01			
15	2	105	0.37	0.74	0.37	0.21	0.00	0.28	0.14	1.56	-0.25	
15	3	105	0.06	1.11	0.06	0.22	0.00	0.30	0.00	0.00	1.15	
15	4	105	0.38	0.74	0.38	0.21	0.00	0.28	0.14	-2.41	6.11	-2.39
15	5	105	0.34	0.78	0.34	0.21	0.00	0.29	0.12	15.02	4.66	1.29
15	6	105	0.38	0.74	0.38	0.21	0.00	0.28	0.14	1.31	2.53	-6.19
15	7	105	0.29	0.85	0.29	0.22	0.00	0.29	0.08	1.21	-680.46	

References

Babin, M., Morel, A., Fournier-Sicre, V., Fell, F., & Stramski, D. (2003). Light scattering properties of marine particles in coastal and open ocean waters as related to the particle mass concentration. *Limnology and oceanography*, *48*(2), 843-859.

Behrenfeld, M.J., Boss, E., Siegel, D.A. and Shea, D.M., (2005). Carbon-based ocean productivity and phytoplankton physiology from space. *Global biogeochemical cycles*, *19*(1).

Boesch, D.F., Brinsfield, R.B. and Magnien, R.E., (2001). Chesapeake Bay eutrophication: Scientific understanding, ecosystem restoration, and challenges for agriculture. *Journal of environmental quality*, *30*(2), pp.303-320.

Boicourt, W. C., Kuzmić, M., & Hopkins, T. S. (1999). The inland sea: circulation of Chesapeake Bay and the Northern Adriatic. *Ecosystems at the land-sea margin: Drainage basin to coastal sea*, *55*, 81-129.

Boynton, W. R., Kemp, W. M., & Keefe, C. W. (1982). A comparative analysis of nutrients and other factors influencing estuarine phytoplankton production. In *Estuarine comparisons* (pp. 69-90). Academic Press.

Boynton, W. R., Garber, J. H., Summers, R., & Kemp, W. M. (1995). Inputs, transformations, and transport of nitrogen and phosphorus in Chesapeake Bay and selected tributaries. *Estuaries*, *18*, 285-314.

Brakebill, J.W., Ator, S.W. and Schwarz, G.E., (2010). Sources of Suspended-Sediment Flux in Streams of the Chesapeake Bay Watershed: A Regional

Application of the SPARROW Model 1. *JAWRA Journal of the American Water Resources Association*, *46*(4), pp.757-776.

Brewin, R. J., Sathyendranath, S., Müller, D., Brockmann, C., Deschamps, P. Y., Devred, E., ... & White III, G. N. (2015). The Ocean Colour Climate Change Initiative: III. A round-robin comparison on in-water bio-optical algorithms. *Remote Sensing of Environment*, *162*, 271-294.

Bronk, D. A., Glibert, P. M., Malone, T. C., Banahan, S., & Sahlsten, E. (1998). Inorganic and organic nitrogen cycling in Chesapeake Bay: autotrophic versus heterotrophic processes and relationships to carbon flux. *Aquatic microbial ecology*, *15*(2), 177-189.

Cao, F., Tzortziou, M., Hu, C., Mannino, A., Fichot, C. G., Del Vecchio, R., ... & Novak, M. (2018). Remote sensing retrievals of colored dissolved organic matter and dissolved organic carbon dynamics in North American estuaries and their margins. *Remote Sensing of Environment*, *205*, 151-165.

Chesapeake Bay Program, (2012). Guide to Using Chesapeake Bay Program Water Quality Monitoring Data. Chesapeake Bay Program, Annapolis, MD. https://d18lev1ok5leia.cloudfront.net/chesapeakebay/documents/wq_data_userguide_10feb12_mod.pdf

Chesapeake Bay Program. (2014). Chesapeake Watershed Agreement. Chesapeake Bay Program, Annapolis, MD. https://www.chesapeakebay.net/what/what-guides-us/watershed-agreement

Ciavatta, S., Torres, R., Martinez-Vicente, V., Smyth, T., Dall'Olmo, G., Polimene, L., & Allen, J. I. (2014). Assimilation of remotely-sensed optical properties to improve marine biogeochemistry modelling. *Progress in Oceanography*, *127*, 74-95.

Cloern, J. E. (1996). Phytoplankton bloom dynamics in coastal ecosystems: a review with some general lessons from sustained investigation of San Francisco Bay, California. *Reviews of Geophysics*, *34*(2), 127-168.

Cloern, J. E., & Jassby, A. D. (2008). Complex seasonal patterns of primary producers at the land–sea interface. *Ecology Letters*, *11*(12), 1294-1303.

Cloern, J. E., & Jassby, A. D. (2010). Patterns and scales of phytoplankton variability in estuarine–coastal ecosystems. Estuaries and coasts, 33, 230-241.

Clune, J.W., and Capel, P.D., eds., 2021, Nitrogen in the Chesapeake Bay watershed— A century of change, 1950–2050 (ver. 1.1, December 2021): U.S. Geological Survey Circular 1486, 168 p., https://doi.org/10.3133/cir1486

Cooper, S. R., & Brush, G. S. (1993). A 2,500-year history of anoxia and eutrophication in Chesapeake Bay. *Estuaries*, *16*, 617-626.

Dall'Olmo, G., Gitelson, A. A., & Rundquist, D. C. (2003). Towards a unified approach for remote estimation of chlorophyll-a in both terrestrial vegetation and turbid productive waters. *Geophysical Research Letters*, *30*(18).

Doerffer, R., & Schiller, H. (2007). The MERIS Case 2 water algorithm. *International Journal of Remote Sensing*, *28*(3-4), 517-535.

Fisher, T. R., Harding Jr, L. W., Stanley, D. W., & Ward, L. G. (1988). Phytoplankton, nutrients, and turbidity in the Chesapeake, Delaware, and Hudson estuaries. *Estuarine, Coastal and Shelf Science*, *27*(1), 61-93.

Fisher, T. R., Peele, E. R., Ammerman, J. W., & Harding Jr, L. W. (1992). Nutrient limitation of phytoplankton in Chesapeake Bay. *Marine Ecology Progress Series*, 51-63.

Friedrichs, M. A. (2001). A data assimilative marine ecosystem model of the central equatorial Pacific: Numerical twin experiments. *Journal of marine research*, *59*(6), 859-894.

Friedrichs, M. A. M. and D. E. Kaufman (2019). Marine biogeochemical data assimilation. In: Encyclopedia of Ocean Sciences (Third Edition). Ed. by J. K. Cochran, H. J. Bokuniewicz, and P. L. Yager. Third Edition. Oxford: Academic Press, 520–526. doi: 10.1016/B978-0-12-409548-9.11261-8.

Gallegos, C. L., Werdell, P. J., & McClain, C. R. (2011). Long-term changes in light scattering in Chesapeake Bay inferred from Secchi depth, light attenuation, and remote sensing measurements. *Journal of Geophysical Research: Oceans*, *116*(C7).

Geider, R. J., MacIntyre, H. L., & Kana, T. M. (1997). Dynamic model of phytoplankton growth and acclimation: responses of the balanced growth rate and the chlorophyll a: carbon ratio to light, nutrient-limitation and temperature. *Marine Ecology Progress Series*, *148*, 187-200.

Gellis, A.C., Hupp, C.R., Pavich, M.J., Landwehr, J.M., Banks, W.S., Hubbard, B.E., Langland, M.J., Ritchie, J.C. and Reuter, J.M., (2009). *Sources, transport, and storage of sediment at selected sites in the Chesapeake Bay Watershed*. U. S. Geological Survey.

Gitelson, A. A., Schalles, J. F., & Hladik, C. M. (2007). Remote chlorophyll-a retrieval in turbid, productive estuaries: Chesapeake Bay case study. *Remote Sensing of Environment*, *109*(4), 464-472.

Gilerson, A. A., Gitelson, A. A., Zhou, J., Gurlin, D., Moses, W., Ioannou, I., & Ahmed, S. A. (2010). Algorithms for remote estimation of chlorophyll-a in coastal and inland waters using red and near infrared bands. *Optics Express*, *18*(23), 24109-24125.

Goodrich, D.M., Boicourt, W.C., Hamilton, P. and Pritchard, D.W., (1987). Wind-induced destratification in Chesapeake Bay. *Journal of Physical Oceanography*, *17*(12), pp.2232-2240.

Gower, J.F.R., Doerffer, R., Borstad, G.A. (1999). Interpretation of the 685 nm peak in water-leaving radiance spectra in terms of fluorescence, absorption and scattering, and its observation by MERIS. Int. J. Remote Sens. (9), 1771–1786.

Gregg, W. W., Friedrichs, M. A., Robinson, A. R., Rose, K. A., Schlitzer, R., Thompson, K. R., & Doney, S. C. (2009). Skill assessment in ocean biological data assimilation. *Journal of Marine Systems*, *76*(1-2), 16-33.

Guildford, S.J. and Hecky, R.E., (2000). Total nitrogen, total phosphorus, and nutrient limitation in lakes and oceans: is there a common relationship? *Limnology and oceanography*, *45*(6), pp.1213-1223.

Gurlin, D., Gitelson, A. A., & Moses, W. J. (2011). Remote estimation of chl-a concentration in turbid productive waters—Return to a simple two-band NIR-red model? *Remote Sensing of Environment*, *115*(12), 3479-3490.

Hagy, J. D., Boynton, W. R., Keefe, C. W., & Wood, K. V. (2004). Hypoxia in Chesapeake Bay, 1950–2001: long-term change in relation to nutrient loading and river flow. *Estuaries*, *27*, 634-658.

Harding, L. W. (1994). Long-term trends in the distribution of phytoplankton in Chesapeake Bay: roles of light, nutrients and streamflow. *Marine Ecology-Progress Series*, *104*, 267-267.

Harding Jr, L. W., & Perry, E. S. (1997). Long-term increase of phytoplankton biomass in Chesapeake Bay, 1950-1994. *Marine Ecology Progress Series*, *157*, 39-52.

Harding Jr, L. W., Magnuson, A., & Mallonee, M. E. (2005). SeaWiFS retrievals of chlorophyll in Chesapeake Bay and the mid-Atlantic bight. *Estuarine, Coastal and Shelf Science*, *62*(1-2), 75-94.

Harding, L.W., Gallegos, C.L., Perry, E.S., Miller, W.D., Adolf, J.E., Mallonee, M.E. and Paerl, H.W., (2016). Long-term trends of nutrients and phytoplankton in Chesapeake Bay. *Estuaries and Coasts*, *39*, pp.664-681.

Harding, L. W., Mallonee, M. E., Perry, E. S., Miller, W. D., Adolf, J. E., Gallegos, C. L., & Paerl, H. W. (2019). Long-term trends, current status, and transitions of water quality in Chesapeake Bay. *Scientific Reports*, *9*(1), 1-19.

Hood, R. R., Shenk, G. W., Dixon, R. L., Smith, S. M. C., Ball, W. P., Bash, J. O., Batiuk, R. A., Boomer, K., Brady, D. C., Cerco, C., Claggett, P., de Mutsert, K., Easton, Z. M., Elmore, A. J., Friedrichs, M. A. M., Harris, L. A., Ihde, T. F., Lacher, L., Li, L., Linker, L. C., Miller, A., Moriarty, J., Noe, G. B., Onyullo, G., Rose, K., Skalak, K., Tian, R., Veith, T. L., Wainger, L., Weller, D., and Zhang, Y. J. (2021) The Chesapeake Bay Program Modeling System: Overview and Recommendations for Future Development, Ecol. Modelling, 465, 109635

Hou, W., Lee, Z. and Weidemann, A.D., (2007). Why does the Secchi disk disappear? An imaging perspective. *Optics Express*, *15*(6), pp.2791-2802.

IOCCG (2006). Remote Sensing of Inherent Optical Properties: Fundamentals, Tests of Algorithms, and Applications. Lee, Z.-P. (ed.), Reports of the International Ocean-Colour Coordinating Group, No. 5, IOCCG, Dartmouth, Canada.

IOCCG (2020). Synergy between Ocean Colour and Biogeochemical/Ecosystem Models. Dutkiewicz, S. (ed.), IOCCG Report Series, No. 19, International Ocean Colour Coordinating Group, Dartmouth, Canada.

Jones, E. M. et al. (2016). Use of remote-sensing reflectance to constrain a data assimilating marine biogeochemical model of the Great Barrier Reef. Biogeosciences 13:23, 6441–6469

Kemp, W. M., Boynton, W. R., Adolf, J. E., Boesch, D. F., Boicourt, W. C., Brush, G., Cornwell, J.C., Fisher, T.R., Glibert, P. M., Hagy, J.D. and Harding, L.W., Houde, E.D., Kimmel, D.G., Miller, W.D., Newell, R.I.E., Roman, M. R., Smith, E. M., Stevenson, J. C. (2005). Eutrophication of Chesapeake Bay: historical trends and ecological interactions. *Marine ecology progress series*, *303*, 1-29.

Knebel, H. J., Martin, E. A., Glenn, J. L., & Needell, S. W. (1981). Sedimentary framework of the Potomac River estuary, Maryland. *Geological Society of America Bulletin*, *92*(8), 578-589.

Langland, M. J., & Cronin, T. M. (Eds.). (2003). *A summary report of sediment processes in Chesapeake Bay and watershed*. US Department of the Interior, US Geological Survey.

Lee, Z., Carder, K. L., & Arnone, R. A. (2002). Deriving inherent optical properties from water color: a multiband quasi-analytical algorithm for optically deep waters. *Applied optics*, *41*(27), 5755-5772.

Malone, T. C., Crocker, L. H., Pike, S. E., & Wendler, B. W. (1988). Influences of river flow on the dynamics of phytoplankton production in a partially stratified estuary. *Marine ecology progress series. Oldendorf*, *48*(3), 235-249.

Malone, T.C., (1991). River flow, phytoplankton production and oxygen depletion in Chesapeake Bay. *Geological Society, London, Special Publications*, *58*(1), pp.83-93.

Mélin, F., & Vantrepotte, V. (2015). How optically diverse is the coastal ocean? *Remote Sensing of Environment*, *160*, 235-251.

Mishra, S., & Mishra, D. R. (2012). Normalized difference chlorophyll index: A novel model for remote estimation of chlorophyll-a concentration in turbid productive waters. *Remote Sensing of Environment*, *117*, 394-406.

Monolisha, S., Platt, T., Sathyendranath, S., Jayasankar, J., George, G., & Jackson, T. (2018). Optical classification of the coastal waters of the Northern Indian Ocean. *Frontiers in Marine Science*, *5*, 87.

Moore, T. S., Dowell, M. D., Bradt, S., & Verdu, A. R. (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.

Moses, W. J., Gitelson, A. A., Berdnikov, S., & Povazhnyy, V. (2009). Satellite estimation of chlorophyll-$ a $ concentration using the red and NIR bands of MERIS—the Azov sea case study. *IEEE Geoscience and Remote Sensing Letters*, *6*(4), 845-849.

Murphy, R. R., Keisman, J., Harcum, J., Karrh, R. R., Lane, M., Perry, E. S., & Zhang, Q. (2021). Nutrient improvements in chesapeake bay: direct effect of load reductions and implications for coastal management. *Environmental Science & Technology*, *56*(1), 260-270.

Nichols, M. M., & Thompson, G. (1973) Development of the turbidity maximum in a coastal plain estuary: Final Report. Virginia Institute of Marine Science, William & Mary. http://dx.doi.org/doi:10.21220/m2-xz1kte69

NOAA National Centers for Environmental Information. (2017). NOAA NOS Estuarine Bathymetry - Chesapeake Bay (M130). National Centers for Environmental Information, NOAA. doi:10.7289/V5ZK5F0X.

Noe, G.B., Cashman, M.J., Skalak, K., Gellis, A., Hopkins, K.G., Moyer, D., Webber, J., Benthem, A., Maloney, K., Brakebill, J. and Sekellick, A., (2020). Sediment dynamics and implications for management: State of the science from long-term research in the Chesapeake Bay watershed, USA. *Wiley Interdisciplinary Reviews: Water*, *7*(4), p.e1454.

O'Reilly, J. E., Maritorena, S., Mitchell, B. G., Siegel, D. A., Carder, K. L., Garver, S. A., ... & McClain, C. (1998). Ocean color chlorophyll algorithms for SeaWiFS. *Journal of Geophysical Research: Oceans*, *103*(C11), 24937-24953.

Pahlevan, N., Smith, B., Schalles, J., Binding, C., Cao, Z., Ma, R., R., Alikas, K., Kangro, K., Gurlin, D., Hà, N. and Matsushita, B., Moses, W., Greb, S., Lehmann, M., Ondrusek, M., Oppelt, N, Stumpf, R. (2020). Seamless retrievals of chlorophyll-a from Sentinel-2 (MSI) and Sentinel-3 (OLCI) in inland and coastal waters: A machine-learning approach. *Remote Sensing of Environment*, *240*, 111604.

Preisendorfer, R. W. (1986). Eyeball optics of natural waters: Secchi disk science.

Pritchard, D.W., (1952). Salinity distribution and circulation in the Chesapeake Bay estuarine system. 1. *Mar. Res*, *11*, pp.106-123.

Redfield, A. C. (1963). The influence of organisms on the composition of seawater. *The sea*, *2*, 26-77.

Roesler, C. S., & Perry, M. J. (1995). In situ phytoplankton absorption, fluorescence emission, and particulate backscattering spectra determined from reflectance. *Journal of Geophysical Research: Oceans*, *100*(C7), 13279-13294.

Roman, M., Zhang, X., McGilliard, C., & Boicourt, W. (2005). Seasonal and annual variability in the spatial patterns of plankton biomass in Chesapeake Bay. *Limnology and Oceanography*, *50*(2), 480-492.

Russ, E., & Palinkas, C. (2020). Evolving sediment dynamics due to anthropogenic processes in upper Chesapeake Bay. *Estuarine, Coastal and Shelf Science*, *235*, 106596.

Sanford, L. P., Suttles, S. E., & Halka, J. P. (2001). Reconsidering the physics of the Chesapeake Bay estuarine turbidity maximum. *Estuaries*, *24*, 655-669.

Sathyendranath, S., Platt, T., Kovač, Ž., Dingle, J., Jackson, T., Brewin, R.J., Franks, P., Marañón, E., Kulk, G. and Bouman, H.A., 2020. Reconciling models of primary production and photoacclimation. *Applied Optics*, *59*(10), pp.C100-C114.

Schubel, J.R., (1968). *Suspended sediment of the northern Chesapeake Bay*. Johns Hopkins Univ Baltimore Md Chesapeake Bay Institute.

Schubel, J.R. and Pritchard, D.W., (1986). Responses of upper Chesapeake Bay to variations in discharge of the Susquehanna River. *Estuaries*, *9*, pp.236-249.

Scott, D. W. (1992), Multivariate Density Estimation: Theory, Practice, and Visualization, New York: John Wiley

Seegers, B. N., Stumpf, R. P., Schaeffer, B. A., Loftin, K. A., & Werdell, P. J. (2018). Performance metrics for the assessment of satellite data products: an ocean color case study. *Optics express*, *26*(6), 7404-7422.

Spyrakos, E., O'donnell, R., Hunter, P. D., Miller, C., Scott, M., Simis, S. G., ... & Tyler, A. N. (2018). Optical types of inland and coastal waters. *Limnology and Oceanography*, *63*(2), 846-870.

Stramski, D., & Woźniak, S. B. (2005). On the role of colloidal particles in light scattering in the ocean. *Limnology and Oceanography*, *50*(5), 1581-1591.

Tango, P. J., & Batiuk, R. A. (2013). Deriving Chesapeake Bay water quality standards. *JAWRA Journal of the American Water Resources Association*, *49*(5), 1007-1024.

Testa, J. M., Lyubchich, V., & Zhang, Q. (2019). Patterns and trends in Secchi disk depth over three decades in the Chesapeake Bay estuarine complex. *Estuaries and Coasts*, *42*(4), 927-943.

Turner, J. S., St-Laurent, P., Friedrichs, M. A., & Friedrichs, C. T. (2021). Effects of reduced shoreline erosion on Chesapeake Bay water clarity. *Science of the Total Environment*, *769*, 145157.

Turner, J. S., Friedrichs, C. T., & Friedrichs, M. A. (2021). Long-Term Trends in Chesapeake Bay Remote Sensing Reflectance: Implications for Water Clarity. *Journal of Geophysical Research: Oceans*, *126*(12), e2021JC017959.

Turner, J. S., Fall, K. A., & Friedrichs, C. T. (2022). Clarifying water clarity: A call to use metrics best suited to corresponding research and management goals in aquatic ecosystems. *Limnology and Oceanography Letters*.

Tzortziou, M., Herman, J. R., Gallegos, C. L., Neale, P. J., Subramaniam, A., Harding Jr, L. W., & Ahmad, Z. (2006). Bio-optics of the Chesapeake Bay from measurements and radiative transfer closure. *Estuarine, Coastal and Shelf Science*, *68*(1-2), 348-362.

Tzortziou, M., Subramaniam, A., Herman, J. R., Gallegos, C. L., Neale, P. J., & Harding Jr, L. W. (2007). Remote sensing reflectance and inherent optical properties in the mid Chesapeake Bay. *Estuarine, Coastal and Shelf Science*, *72*(1-2), 16-32.

U.S. Environmental Protection Agency. (2010). Chesapeake Bay Total Maximum Daily Load for Nitrogen, Phosphorus, and Sediment. U.S. Environmental Protection Agency, Annapolis, MD. https://www.epa.gov/chesapeake-bay-tmdl/chesapeake-bay-tmdl-document

Vantrepotte, V., Loisel, H., Dessailly, D., & Mériaux, X. (2012). Optical classification of contrasted coastal waters. *Remote Sensing of Environment*, *123*, 306-323.

Werdell, P.J., McKinna, L.I., Boss, E., Ackleson, S.G., Craig, S.E., Gregg, W.W., Lee, Z., Maritorena, S., Roesler, C.S., Rousseaux, C.S. and Stramski, D., (2018). An overview of approaches and challenges for retrieving marine

inherent optical properties from ocean color remote sensing. *Progress in oceanography*, *160*, pp.186-212.
Werdell, P. J., Behrenfeld, M. J., Bontempi, P. S., Boss, E., Cairns, B., Davis, G. T., Franz, B.A., Gliese, U.B., Gorman, E.T., Hasekamp, O. and Knobelspiesse, K.D, Mannino, A., Martins, J.V., McClain, C.R., Meister, G., & Remer, L. A. (2019). The Plankton, Aerosol, Cloud, ocean Ecosystem mission: status, science, advances. *Bulletin of the American Meteorological Society*, *100*(9), 1775-1794.
Windle, A. E., Evers-King, H., Loveday, B. R., Ondrusek, M., & Silsbe, G. M. (2022). Evaluating atmospheric correction algorithms applied to olci sentinel-3 data of Chesapeake Bay Waters. *Remote Sensing*, *14*(8), 1881.
Wolny, J. L., Tomlinson, M. C., Schollaert Uz, S., Egerton, T. A., McKay, J. R., Meredith, A., Reece, K.S., Scott, G.P., & Stumpf, R. P. 2020. Current and future remote sensing of harmful algal blooms in the Chesapeake Bay to support the shellfish industry. *Frontiers in Marine Science*, *7*, 337.
Xu, J., & Hood, R. R. (2006). Modeling biogeochemical cycles in Chesapeake Bay with a coupled physical–biological model. *Estuarine, Coastal and Shelf Science*, *69*(1-2), 19-46.
Zawada, D. G., Hu, C., Clayton, T., Chen, Z., Brock, J. C., & Muller-Karger, F. E. (2007). Remote sensing of particle backscattering in Chesapeake Bay: A 6-year SeaWiFS retrospective view. *Estuarine, Coastal and Shelf Science*, *73*(3-4), 792-806.
Zhang, Q., Brady, D. C., Boynton, W. R., & Ball, W. P. (2015). Long-term trends of nutrients and sediment from the nontidal Chesapeake watershed: An assessment of progress by river and season. *JAWRA Journal of the American Water Resources Association*, *51*(6), 1534-1555.
Zhang, Q., & Blomquist, J. D. (2018). Watershed export of fine sediment, organic carbon, and chlorophyll-a to Chesapeake Bay: Spatial and temporal patterns in 1984–2016. *Science of the Total Environment*, *619*, 1066-1078.
Zhang, Q., Hirsch, R. M., & Ball, W. P. (2016). Long-term changes in sediment and nutrient delivery from Conowingo Dam to Chesapeake Bay: Effects of reservoir sedimentation. *Environmental Science & Technology*, *50*(4), 1877-1886.
Zhang, Q., Fisher, T. R., Trentacoste, E. M., Buchanan, C., Gustafson, A. B., Karrh, R., ... & Tango, P. J. (2021). Nutrient limitation of phytoplankton in Chesapeake Bay: Development of an empirical approach for water-quality management. *Water Research*, *188*, 116407.
Zheng, G., & DiGiacomo, P. M. (2017). Remote sensing of chlorophyll-a in coastal waters based on the light absorption coefficient of phytoplankton. *Remote Sensing of Environment*, *201*, 331-341.

Chapter 4: Evaluation of Unoccupied Aircraft System (UAS) Remote Sensing Reflectance Retrievals for Water Quality Monitoring in Coastal Waters

Introduction

Unoccupied aircraft systems (UAS, drones) provide on-demand remote sensing capabilities at ultra-high resolution (<5 cm) without the challenges of cloud cover, land adjacency, and atmospheric effects associated with satellite and airborne remote sensing (Anderson & Gaston, 2013). UAS are becoming an integral tool in studying and managing coastal ecosystems (Johnston, 2019), with applications ranging from thermal remote sensing (Lee et al., 2016; Dugdale et al., 2019) and pollutant tracking (Arango & Nairn, 2019; Morgan et al., 2020), to habitat and population assessments (Gray et al., 2018; Windle et al., 2019). The increased spatial and temporal resolution provided by UAS remote sensing can enhance ecological and biogeochemical research of aquatic ecosystems. UAS has the potential to characterize the degree of eutrophication, identify the extent and movement of harmful algal blooms, and resolve fine-scale coupled biophysical processes in coastal and inland water bodies.

A number of recent studies have derived optical water quality parameters using UAS imagery and are listed in Table 1. In these studies, UAS are equipped with a variety of either multispectral or hyperspectral imagers that measure light at discrete wavebands in the visible and near-infrared (NIR) spectrum. Following the large body of research borne from earth observing satellites (Werdell & McClain, 2019), many of these studies use a combination of UAS optical payloads, calibrations, and numerical methods to determine remote sensing reflectance (R_{rs}) defined as:

$$R_{rs}(\theta, \phi, \lambda) = \frac{L_W(\theta,\phi,\lambda)}{E_d(\lambda)} \tag{1}$$

where L_W (W m^{-2} nm^{-1} sr^{-1}) is water-leaving radiance, E_d (W m^{-2} nm^{-1}) is downwelling irradiance, θ represents the sensor viewing angle between the sun and the vertical (zenith), ϕ represents the angular direction relative to the sun (azimuth), and λ represents wavelength.

Like all above-water optical measurements, UAS do not measure R_{rs} directly as the at-sensor total radiance (L_T, W m^{-2} nm^{-1} sr^{-1}) constitutes the sum of L_W and incident radiance reflected off the sea surface into the detector's field of view, herein referred to as surface-reflected radiance (L_{SR}). L_W is radiance that emanates from the water and contains a spectral shape and magnitude governed by optically active water constituents interacting with downwelling irradiance, while L_{SR} is independent of water constituents and is instead governed by a given sea-state surface reflecting light; a familiar example is sun glint. Here we define UAS total reflectance (R_{UAS}) as:

$$R_{UAS}(\theta, \phi, \lambda) = \frac{L_T(\theta,\phi,\lambda)}{E_d(\lambda)} \tag{2}$$

where

$$L_T(\theta, \phi, \lambda) = L_W(\theta, \phi, \lambda) + L_{SR}(\theta, \phi, \lambda) \tag{3}$$

As UAS measurements are typically performed close to the surface (e.g., United States Federal Aviation Administration's maximum allowable altitude of 122 meters), atmospheric measurement effects are routinely assumed to be negligible and ignored (Zeng et al., 2017; Schneider-Zapp, 2019). Indeed, this is a key advantage of UAS imagery as atmospheric effects over coastal environments can introduce significant uncertainty in satellite-based measurements (Gordon & Clark, 1980). However, failure to accurately account for L_{SR} can significantly influence R_{rs}, resulting in inaccurate water quality estimates once algorithms are applied (Zeng et al., 2017; Su, 2017).

The objective of this paper is to evaluate the efficacy of methods that remove L_{SR} from total UAS radiance measurements in order to derive more accurate remotely sensed retrievals of scientifically valuable in-water constituents. UAS-derived radiometric measurements are evaluated against *in situ* hyperspectral R_{rs} measurements to determine the best performing method of estimating and removing surface reflected light and derived water quality estimates.

Table 1. Summary of existing UAS aquatic remote sensing literature including the UAS sensor(s) used, radiometric quantity studied (where R_{rs} represents UAS-derived remote sensing reflectance and R_{UAS} represents UAS-derived total reflectance), whether the study accounted for surface reflected radiance (L_{SR}), and the water quality parameter(s) derived.

Reference	**UAS sensor(s)**	**Radiometric quantity**	**Removal of L_{SR}?**	**WQ parameter(s)**
Zeng et al., 2017	Ocean Optics STS-VIS spectrometers (hyperspectral)	R_{UAS}	No	Chl *a*, CDOM, turbidity

Shang et al., 2017	AvaSpec-dual spectroradiometers (hyperspectral)	R_{rs}	Yes	Chl *a*
Su, 2017	Canon Powershot S110 RGB and NIR sensors	R_{UAS}	Yes	Chl *a,* secchi disk depth, turbidity
Choo et al., 2018	MicaSense RedEdge and DLS (multispectral)	R_{UAS}	No	Chl *a*
Baek et al., 2019	MicaSense RedEdge and DLS (multispectral)	R_{rs}	Yes	Chl *a*
Becker et al., 2019	Ocean Optics STS-VIS spectrometers (hyperspectral)	R_{UAS}	No	Cyanobacteria index, Chl *a* TSS
Arango et al., 2019	MicaSense RedEdge and DLS (multispectral)	R_{UAS}	No	Secchi disk depth, Chl *a,* TSS, TN, TP
Olivetti et al., 2020	Parrot Sequoia (multispectral)	R_{UAS}	No	TSS
McEliece et al., 2020	Sentera multispectral sensor (4 visible bands)	R_{UAS}	No	Chl *a*, turbidity
Kim et al., 2020	MicaSense RedEdge-M and DLS (multispectral)	R_{rs}	Yes	Chl *a* (but not focus of paper)
Castro et al., 2020	MicaSense RedEdge and DLS (multispectral)	R_{rs}	No	Chl *a*
O'Shea et al., 2020	Resonon Pika L spectrometer (hyperspectral) *deployed on a tower, not UAS	R_{rs}	Yes	Chl *a*

Background/Theory

If a water surface was perfectly flat, incident light would reflect specularly and could be measured with known viewing geometries. This specular reflection of a level surface is known as the Fresnel reflection; however, most water bodies are not flat as winds and currents create tilting surface wave facets. Due to differing orientation of wave facets reflecting radiance from different parts of the sky, L_{SR} can vary widely within a single image. A common approach to model L_{SR} is to express it as the product of sky radiance (L_{sky}, W m^{-2} nm^{-1} sr^{-1}) and ρ, the effective sea-surface reflectance of the wave facet (Mobley, 1999; Lee et al., 2010):

$$L_{SR}(\theta, \phi, \lambda) = \rho(\theta, \phi, \lambda) * L_{sky}(\theta, \phi, \lambda) \qquad (4)$$

Rearranging equations 3 and 4, ρ can be derived by:

$$\rho(\theta, \phi, \lambda) = \frac{L_T(\theta,\phi,\lambda) - L_w(\theta,\phi,\lambda)}{L_{sky}(\theta,\phi,\lambda)} \qquad (5)$$

Given measurements of L_{sky}, an accurate determination of ρ is critical to derive R_{rs} by:

$$R_{rs}(\theta, \phi, \lambda) = R_{UAS}(\theta, \phi, \lambda) - \frac{L_{sky}(\theta,\phi,\lambda) * \rho(\theta,\phi,\lambda)}{E_d(\lambda)} \qquad (6)$$

Methods using the statistics of the sea surface, for a given wind vector, can be used to predict ρ; tabulated values have been derived from numerical simulations with modelled sea surfaces, Cox and Munk wave states (wind), and viewing geometries (Cox & Munk, 1954; Mobley, 1999; Mobley, 2015). Mobley (1999) provides the recommendation of collecting radiance measurements at viewing directions of θ = 40° from nadir and ϕ = 135° from the sun to minimize the effects of sun glint and nonuniform sky radiance with a ρ value of 0.028. The suggested viewing geometries and ρ value from Mobley (1999) have been used to estimate and remove L_{SR} in UAS remote sensing studies (Shang et al., 2017; Baek et al., 2019; Kim et al., 2020). However, more recent studies have shown that ρ can be spectrally dependent due to the degree of sky polarization and the ratio of diffuse to direct light; assuming a spectally constant ρ in a roughened sea surface with multi-angular wave facets can lead to erroneous estimates of L_{SR} (Lee et al., 2010; Mobley, 2015). Lee et al. (2010) proposed a spectral optimization approach using spectral inherent optical properties to model $R_{rs,}$ which has been applied in UAS remote sensing studies (O'Shea et al., 2020).

An alternative method to remove L_{SR} relies on the so-called dark pixel assumption that assumes L_W in the NIR is negligible due to strong absorption of water. Where this assumption holds, at-sensor radiance measured in the NIR is solely L_{SR} (Gordon & Wang, 1994; Siegel et al., 2000) and allows ρ to be calculated if L_{sky} is known. Studies have used this assumption to estimate and remove L_{SR}; however, the assumption tends to fail in more turbid waters where high concentrations of particles enhance backscattering and L_W in the NIR (Siegel et al., 2000; Lavender et al., 2005).

Other novel image-processing techniques such as nonlocal mean filtering and a matching pixel by pixel algorithm have been proposed to reduce the effects of L_{SR} variation from UAS imagery; however, these region-specific methods exhibit some limitations for applications in other water bodies (Su, 2017; Totsuka et al., 2019).

Methods

Study area

This study was conducted in the Choptank River, a major tributary of the Chesapeake Bay (USA). The 1,756 km^2 coastal plain watershed is dominated by agriculture and forest with a relatively low population density, transitions from non-tidal, freshwater reaches to a brackish, tidal mouth, and is eutrophic (Fisher et al., 2021). Data from eight stations located downriver and extending from the mouth of the Choptank River were collected on September 16, 2020 with a relatively clear sky, data from eight stations located up river were collected on October 1, 2020 with varying cloud cover, and data from thirteen stations located in between the downriver and upriver stations were collected on October 22, 2020 with a clear sky (Figure 1).

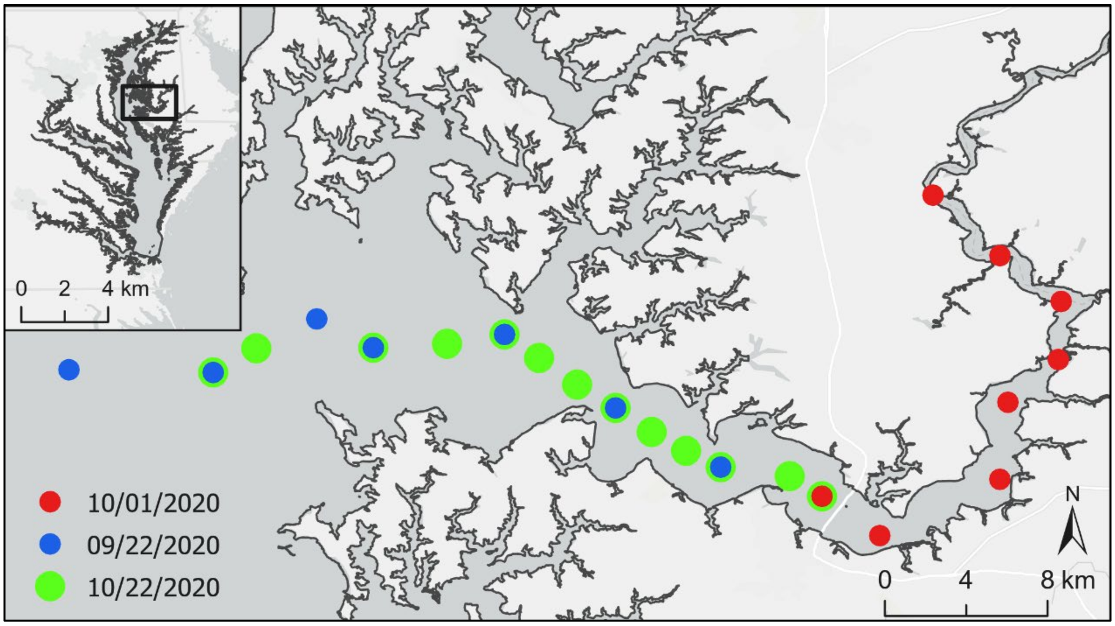


Figure 1. Location of Choptank River (38.63° N, -76.33 W) in Chesapeake Bay, Maryland, United States and locations of stations where data was collected.

In situ Rrs

At each station, a set of hyperspectral radiometers (TriOS RAMSES; Rastede DE) deployed on a float provided *in situ* R_{rs} measurements at every station. The float held a downwelling irradiance sensor (E_d) and an inverted radiance sensor (L_w) positioned above the surface water and with a small black plastic cone that extended just below the surface to block L_{SR} (i.e., light blocking technique, Ahn 1999; Lee et al., 2019). The TriOS radiometers collect 256 wavelength bands at 3.3 nm intervals within the 320–950 nm range. All *in situ* hyperspectral measurements were interpolated at 1 nm intervals. In order to compare these measurements to the UAS, spectral response functions (SRFs) of the five MicaSense wavebands were applied to the *in situ* hyperspectral R_{rs} data (Figure 2).

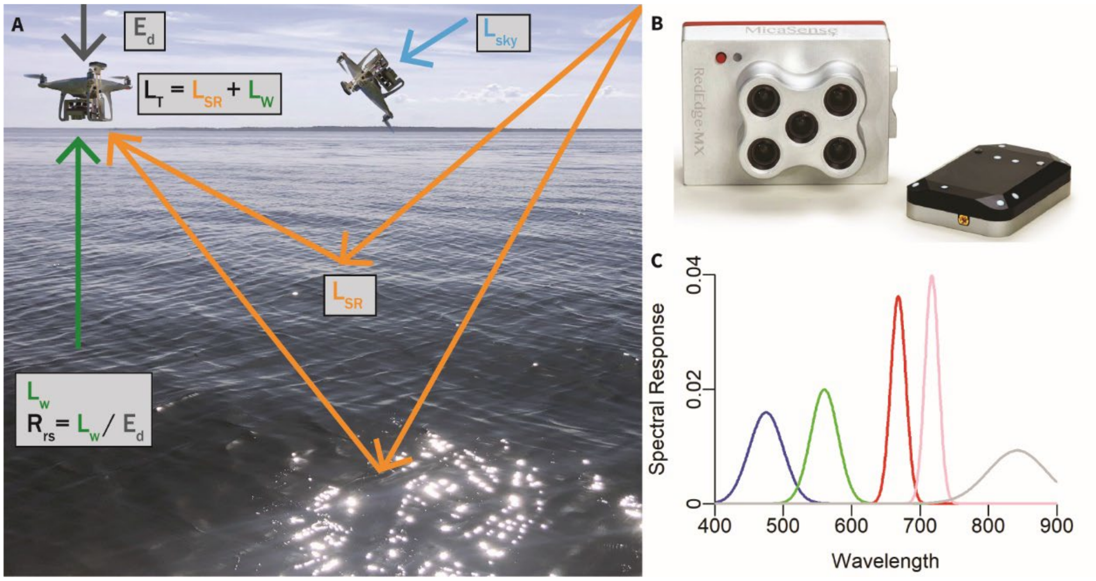


Figure 2. A) MicaSense RedEdge-MX multispectral sensor and downwelling light sensor (DLS) collects total radiance (L_T) (sum of water-leaving radiance, L_w and surface reflected radiance, L_{SR}) and downwelling irradiance (E_d) measurements while in a low altitude flight. Sky radiance (L_{sky}) is collected by positioning the sensor at a 40° angle from zenith away from sun. B) MicaSense RedEdge-MX multispectral sensor and DLS collects L_T and E_d in 5 wavebands: red, green, blue, red-edge, near-infrared. C) Approximate spectral response functions of MicaSense RedEdge-MX sensor.

In situ water quality data

At each station, surface water grab samples were collected to measure chlorophyll *a* and total suspended solids (TSS) concentrations. Chlorophyll *a* concentration was measured in duplicate following EPA method 445.0 (Arar & Collins, 1997). Briefly, 100 ml of water from each station was filtered on 47 mm GF/F filters, immersed in 20 ml of 90% acetone and placed in a dark freezer for 24 hours. Chlorophyll *a* fluorescence was measured on 5 ml of extract using a fluorometer (Turner 10 AU Fluorometer, San Jose CA) calibrated against chlorophyll *a* pigment standards (DHI, Horsholm DK) before and after acidification with 0.1 ml of 6 N hydrochloric acid. TSS was measured using a gravimetric analysis (APHA, 1995). 350 ml of water from each station was filtered using pre-weighed and combusted (450° C) GF/F filters and placed into a 105° C drying oven for at least two hours. Filters were reweighed and the concentration was calculated by TSS (mg/L) = $W_{post(g)}$ - $W_{combust(g)}$ x 1000 / V(L).

Multispectral sensor: MicaSense RedEdge-MX sensor

The MicaSense RedEdge-MX sensor (MicaSense, Seattle, Washington, USA) is a 8.7 x 5.9 x 4.5 cm multispectral camera capable of capturing five simultaneous bands on the electromagnetic spectrum in 12 bit radiometric resolution: blue (475 nm center, 32 nm bandwidth), green (560 nm center, 27 nm bandwidth), red (668 nm center, 14 nm bandwidth) , red edge (717 nm center, 12 nm bandwidth), and NIR (842 nm center, 57 nm bandwidth). The sensor has up to one capture per second, a 47.2° field of view with a spatial resolution of 8 cm per pixel at 120 m altitude. The sensor was mounted on a DJI Phantom 4 Pro UAS using a 10° 3D-printed mount which resulted in a direct nadir viewing angle while in flight (Figure 2). The sensor also includes a downwelling light sensor (DLS) which measures E_d in the same spectral wavebands during in-flight image captures. The DLS measures light incident on a diffuser, providing a downwelling hemispherical irradiance measurement (Mamaghani & Salvaggio, 2019). The DLS was mounted above the UAS to eliminate shading and collected incident light at a 0° zenith angle. Images were collected at an average flight altitude of 70 m which resulted in 0.02m pixel resolution (923x1219).

At each station, measurements of L_T, L_{sky}, and E_d were collected by the MicaSense RedEdge-MX multispectral sensor and DLS. First, measurements of L_{sky} were collected by positioning the camera 40° from zenith with an approximate azimuthal viewing direction of 135° and taking several image captures (Figure 2). L_{sky} at each waveband was computed as the grand mean of all station-specific measurements. Next, the UAS was manually flown and images were automatically captured every two seconds. The MicaSense multispectral sensor has a built-in global positioning system (GPS) and inertial measurement unit (IMU) which recorded the positioning (latitude, longitude, altitude) and orientation (yaw, pitch, roll) of each image capture. Thirty image captures taken at the highest altitude from each station were used in subsequent analyses. This information along with solar elevation, image size, and coordinated universal time (UTC) time, were recorded in the metadata of each image capture.

Pre-processing

Data collected by the multispectral sensor were radiometrically calibrated using a Python workflow provided by MicaSense[7]. The process of converting raw

---

[7] MicaSense RedEdge Image Processing Tutorials. Retrieved online at https://github.com/micasense/imageprocessing.

pixel values (digital number, DN) into total spectral radiance (L_T) values with units of $W/m^2/sr/nm$ is given in Eq 7:

$$L_T = V(x,y) * \frac{a_1}{g} * \frac{DN - DN_{BL}}{t_e + a_2 y - a_3 t_e y} \quad (7)$$

The radiometric calibration compensates for dark pixel subtraction (DN_{BL}), sensor gain (g), exposure settings (t_e), and lens vignette effects (V(x,y)). Coefficients a_1, a_2, and a_3 are radiometric calibration coefficients and x,y is the pixel column and row number, respectively. Lens distortion effects, such as band-to-band image alignment, were removed from image captures by an unwarping procedure and wavebands were aligned to form a stacked TIFF for each image set.

A filtering procedure was applied to all images to remove specular sun glint and instances of the boat when it was present in the image. Studies have suggested filtering out high total radiance values, or instances of specular sun glint, before removing L_{SR} from L_T (Hooker et al., 2002). Specular sun glint arises when direct sunlight reflects off of a wave facet or surface at the viewing angle of the sensor. Sun glint pixels were masked using an empirical upper limit of R_{UAS} measurements in the NIR, and this mask was then propagated to pixels at all other wavebands. Specifically, *in situ* measurements and values from radiative transfer model simulations (see methods below) were used in Eq. 8 to solve for an upper limit of R_{UAS}(NIR) and pixels with values greater than this upper limit were masked in each waveband.

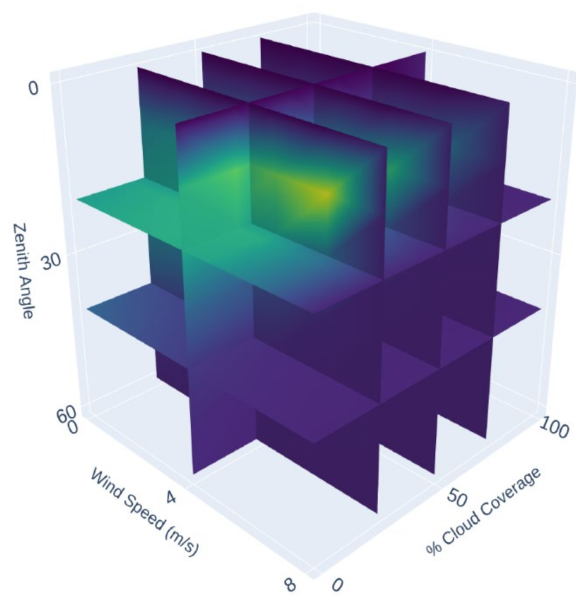
$$R_{UAS}(NIR) = R_{rs}(NIR) + \rho(\mathrm{NIR}) * \frac{L_{sky}}{E_d} \quad (8)$$

Where $R_{rs}(NIR) = 0.005$ is based on *in situ* measurements in this study and consistent with other turbid waters (Tzortziou et al., 2006), ρ(NIR) varied depending on station, and $L_{sky}/E_d = 0.39$ is derived from radiative transfer model simulations (HydroLight v6.0, Numerical Optics Ltd, U.K.). A lower limit of $R_{UAS}(green) = 0.007$ was used to mask out the dark canopy of the boat when present in images.

Removal of surface reflected light (L_{SR})

At each station, L_{SR} was estimated using one of four different methods to ultimately derive R_{rs} following Equation 6. These methods are provided below, and herein referred to as 'ρ_{LUT}', 'NIR=0', 'NIR>0', and 'Deglinting'. Resultant UAS R_{rs} estimates were then compared to paired *in situ* R_{rs} data across stations and wavebands. Statistical evaluations to assess the relationship included root mean square error (RMSE), relative root mean squared error (RRMSE), coefficient of determination (R^2), and p-values.

ρ_{LUT}. The ρ_{LUT} method follows from Mobley (1999) and involved developing a look-up table (LUT) of ρ values using HydroLight simulations. HydroLight is a numerical model that solves the radiative transfer equation to compute the radiance distribution within and at the surface of a water body. Inputs include absorbing and scattering properties of a water body, the nature of the wind-blown sea surface (Cox-Munk sea surface slopes), and the sun and sky radiance incident on the sea surface (Mobley, 1999). Outputs include the full radiance distribution, including an effective ρ value for a range of geometries and wavelengths. Batch HydroLight runs incorporating varying solar zenith angles (0, 10, 20, 30, 40, 50, 60°), wind speeds (0, 2, 4, 6, 8 m/s), and fractional cloud cover (0, 0.2, 0.4, 0.6, 0.8, 1) were used to develop a multidimensional look-up table of ρ values that correspond to a nadir viewing angle (Figure 3). HydroLight returned spectrally explicit ρ(θ, ϕ, λ) values and MicaSense SRFs were applied to ρ(θ, ϕ, λ) for each waveband. ρ(θ, ϕ, λ) values were obtained for each station according to the solar zenith angles obtained from MicaSense metadata, wind speed was collected from the Global Forecast System[8], and cloud cover determined by the L_{sky} images. Data in this study were collected at solar zenith angles ranging from 36.5 to 58.6°, wind speeds ranged from 1.1 to 3.1 m/s and cloud cover ranged from 0 to 60%.



[8] Wind speed data was retrieved from https://earth.nullschool.net/about.html

Figure 3. Visualization of the three-dimensional ρ look-up table derived from HydroLight simulations with varying solar zenith angles, wind speed (m/s), and cloud cover (%) corresponding to a nadir viewing angle.

NIR=0. L_{SR} was also estimated using a pixel based dark pixel assumption to derive ρ. Assuming R_{rs}(NIR) equals 0, Eq. 6 can be rearranged to solve for ρ (Eqs 9-10). This ρ value is used to calculate R_{rs} across all wavebands (Eq 11).

$$0 = L_T(NIR) - L_{sky}(NIR) * \rho \qquad (9)$$

$$\rho = L_T(NIR)/L_{sky}(NIR) \qquad (10)$$

$$R_{rs}(\theta, \phi, \lambda) = R_{UAS}(\theta, \phi, \lambda) - \frac{L_{sky}(\theta,\phi,\lambda) * (L_T(NIR)/L_{sky}(NIR))}{E_d(\lambda)} \qquad (11)$$

NIR > 0. Since the dark pixel assumption is invalid in turbid waters including the Chesapeake Bay (Siegel et al., 2000) an alternative pixel based approach was developed to instead estimate a baseline R_{rs}(NIR). Specifically, R_{rs}(NIR) was estimated empirically by a nonlinear regression using *in situ* R_{rs}(NIR) and R_{UAS} data (Figure 4). The ratio of R_{UAS}(blue)/R_{UAS}(rededge) led to the most robust predictions of R_{rs}(NIR) using Eq 12:

$$R_{rs}(\theta, \phi, \lambda) = 0.025e^{-5.469 * R_{UAS}(blue)/R_{UAS}(red\ edge)} + 0.00013 \qquad (12)$$

Conceptually, this relationship makes sense because with increasing particle concentrations, water becomes less blue and R_{rs}(NIR) increases.

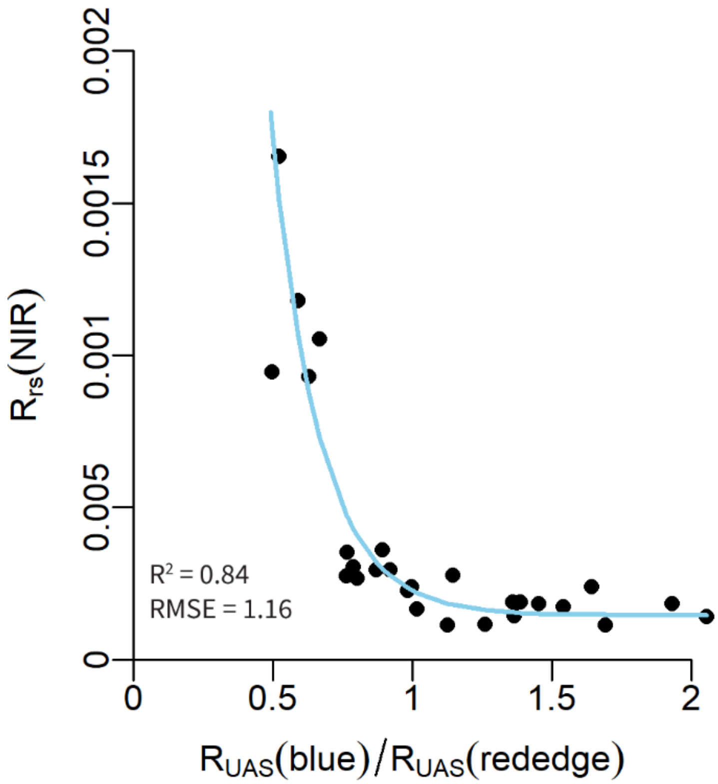


Figure 4. Nonlinear relationship between a blue (475 nm) to red edge (717 nm) ratio of total UAS reflectance to *in situ* R_{rs} in the NIR band (842 nm). An exponential model was used to estimate UAS derived R_{rs}(NIR) baseline in order to estimate L_{SR}.

'Deglinting'. L_{SR} was also estimated following the 'deglinting' methods of Hochberg et al. (2003) and Hedley et al. (2005). For each station, a minimum NIR value was determined by finding the lowest 10% of R_{UAS}(NIR) across all images. For each band, a linear regression was made between all R_{UAS}(NIR) and R_{UAS}(visible) values and the slope (b_i) was determined. Each pixel was corrected by subtracting the product of b_i and the NIR brightness of the pixel (Hedley et al., 2005):

$$R_{rs}(i) = R_{UAS}(i) - b_i(R_{UAS}(NIR) - \min(R_{UAS}(NIR))) \quad (13)$$

Water quality retrievals

R_{rs} values from each L_{SR} removal method were compared against *in situ* chlorophyll *a* and TSS data (n=28) using multiple linear regressions. The best performing model (highest R^2 and lowest RMSE, RRMSE, p-values) was used in a stepwise model selection by Akaike information criterion (AIC, 'stepAIC' in R). The AIC stepwise regression iteratively added and removed wavelengths in order to determine the combination of data that resulted in the best performing model with low prediction error, while taking into account model simplicity. R_{rs} values were used as input into optical algorithms derived from the best performing multiple linear regressions and mean chlorophyll *a* and TSS concentration at each station was obtained by averaging values across all images. The resulting arrays were georeferenced using the Python libraries 'CameraTransform' (Gerum et al., 2019) and 'Rasterio' using archived metadata including latitude, longitude, image width, image height to position the images accurately in a known coordinate system (WGS84). Georeferenced arrays were exported as individual TIFFs and mapped using ArcGIS Pro (ESRI Inc., Redlands, CA, USA).

Results

In situ data

Chlorophyll *a* concentration ranged from 5.19 to 53.30 μg/L with an average of 17.13 ± 10.96 μg/L. TSS concentration ranged from 19.94 mg/L to 39.69 mg/L with an average of 28.25 ± 5.21 mg/L. In general, *in situ* R_{rs} spectra were representative of a eutrophic system (Figure 5, Gitelson, 2007; Spyrakos et al., 2017). Due to absorption of chromophoric dissolved organic matter (CDOM) and chlorophyll *a* in lower wavelengths, R_{rs} is low in the blue region with a distinct peak in the green region (~550 nm). A secondary peak in the red region (~700 nm) corresponds to chlorophyll fluorescence emission (Spyrakos et al., 2017). Overall, spectra collected from upriver stations on 10/01/20 were larger in magnitude and contained a tertiary peak in the NIR region (~800 nm).

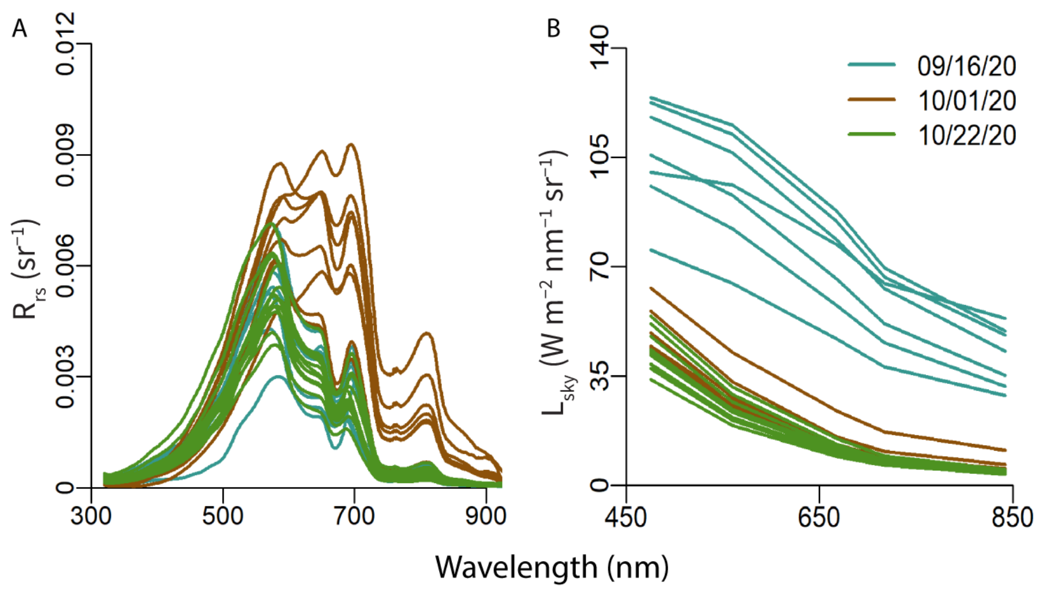


Figure 5. A) *In situ* remote sensing reflectance (R_{rs}) spectra collected with hyperspectral TriOS radiometers and B) Sky radiance (L_{sky}) spectra collected with MicaSense RedEdge-MX multispectral camera at stations along the Choptank River on 9/16/20 (blue), 10/01/20 (brown), and 10/22/20 (green).

Removal of surface reflected light (L_{SR})

UAS measurements of L_{sky} were approximately proportional to the fourth power of the wavelength ($L_{sky} \approx \lambda^{-4}$, Figure 5) and were used with estimations of ρ to calculate and remove L_{SR} from R_{UAS} measurements at each station. Figure 6 illustrates the initial masking and methods to remove L_{SR} using an individual UAS image taken in the green band. When L_{SR} was estimated using the ρ look-up table (ρ_{LUT}), R_{rs} values did not change much from R_{UAS} values; however, R_{rs} values using the pixel-based dark pixel assumptions (NIR=0, NIR>0) and 'deglinting' approach declined and became more homogenous across pixels (Figure 6).

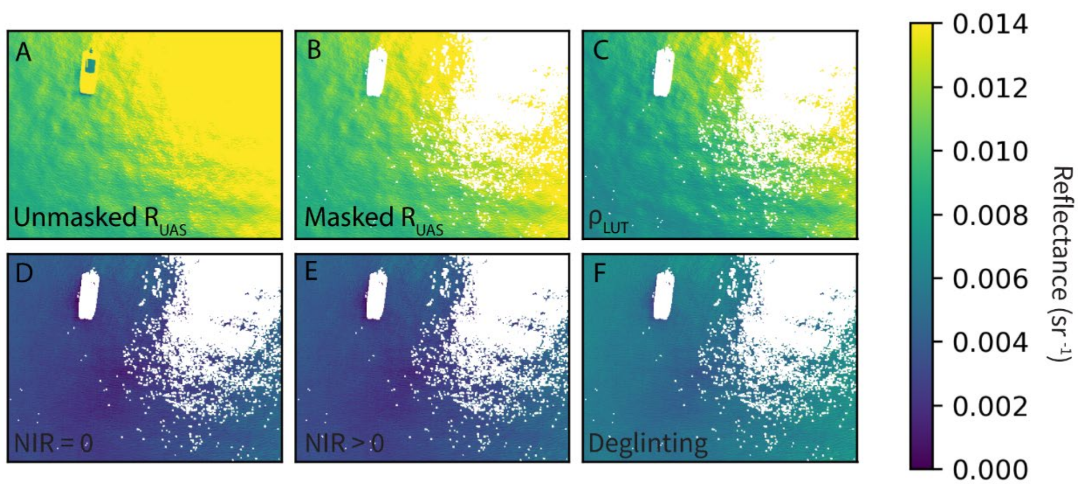


Figure 6. Example of an individual UAS image (green band) with different radiometric values: (A) R_{UAS}, (B) R_{UAS} with initial sun glint masking and (C-F) remote sensing reflectance (R_{rs}) using various

methods to remove surface reflected light: (C) ρ look-up table (LUT), (D) Dark pixel assumption with NIR = 0, E) Dark pixel assumption with NIR > 0, F) Deglingting methods following Hochberg et al. (2003).

UAS-derived reflectance measurements from each L_{SR} removal method were plotted as spectra (Figure 7) and compared against *in situ* R_{rs} values (Figure 8). UAS reflectance spectra are similar in shape to *in situ* R_{rs} spectra and display a distinct peak in the green band (560 nm) and often a secondary peak in the red band (668 nm) corresponding to chlorophyll *a* reflectance and fluorescence, respectively (Figure 7). Spectra collected on the second sampling day (upriver sites) shift to longer wavelengths likely due to scattering of inorganic particles (Figure 7). R_{UAS} spectra are higher in magnitude and are overestimated when compared to *in situ* R_{rs} due to the inclusion of L_{SR} ($R^2 = 0.55$, RMSE = 0.004, $p < 0.05$). When L_{SR} is estimated using the ρ_{LUT}, spectra are similar in shape to *in situ* spectra, but are higher in magnitude and overestimated compared to *in situ* measurements ($R^2 = 0.85$, RMSE = 0.003, $p < 0.05$). When L_{SR} is estimated with the dark pixel assumption (NIR=0), R_{rs} spectra are lowest in magnitude, had the poorest fit to *in situ* data ($R^2 = 0.18$, RMSE = 0.004, $p < 0.05$), and produced negative R_{rs} values in the lower wavelengths. This is not surprising given that the *in situ* measurements in the NIR (Figure 5A) are not negligible which leads to an overestimation in L_{SR}. When L_{SR} is estimated using a basline NIR value (NIR>0), R_{rs} spectra are higher in magnitude with a lower error but still indicate negative values in the lower wavelengths, indicating L_{SR} was still slightly overestimated ($R^2 = 0.50$, RMSE = 0.002, $p < 0.05$). R_{rs} spectra with the

'Deglinting' approach are similar in shape to *in situ* spectra but are slightly higher in magnitude (Figure 7). This approach contained the second highest correlation and lowest RMSE ($R^2 = 0.65$, RMSE = 0.002, $p < 0.05$), but still tended to overestimate R_{rs} values (Figure 8).

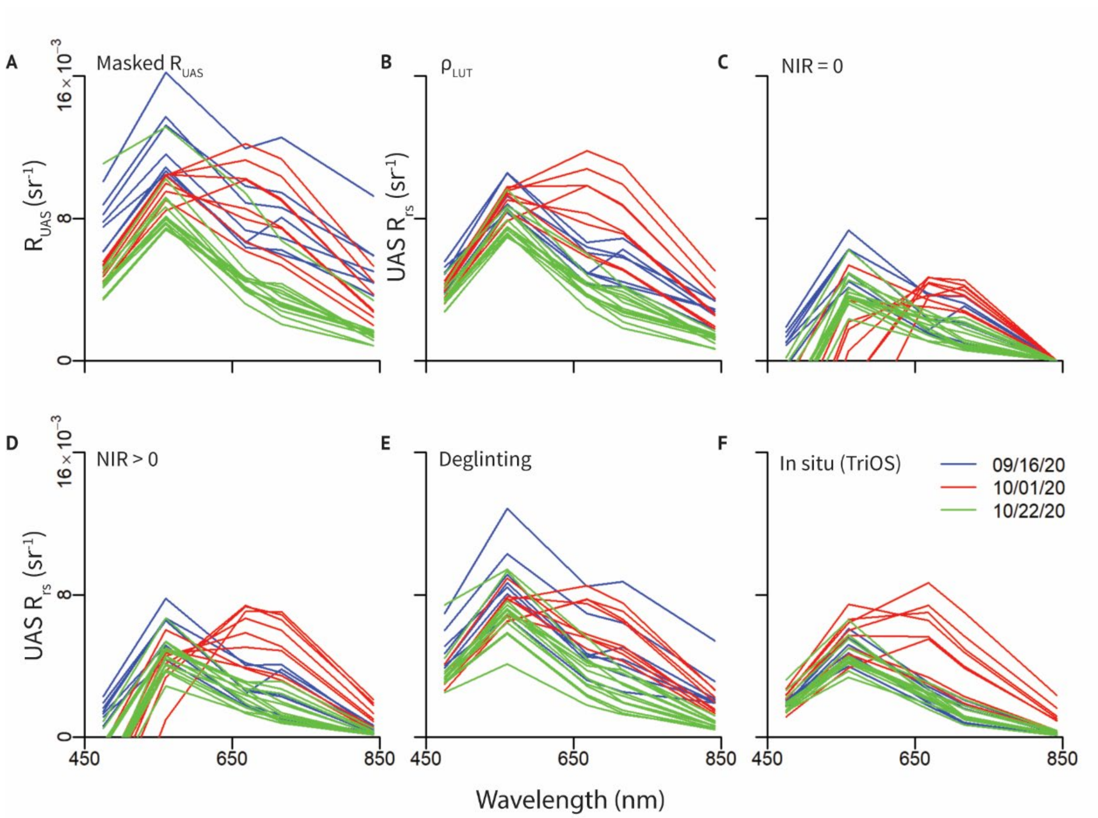


Figure 7. Total UAS reflectance (R_{UAS}) (A) and remote sensing reflectance (R_{rs}) spectra using various methods to remove surface reflected light: B) ρ look-up table from HydroLight simulations, C) Dark pixel assumption with NIR = 0, D) Dark pixel assumption with NIR > 0, E) Deglinting methods following Hochberg et al. (2003), and F) *In situ* R_{rs} spectra from TriOS sensors with MicaSense SRFs applied. Negative values are not shown in plots.

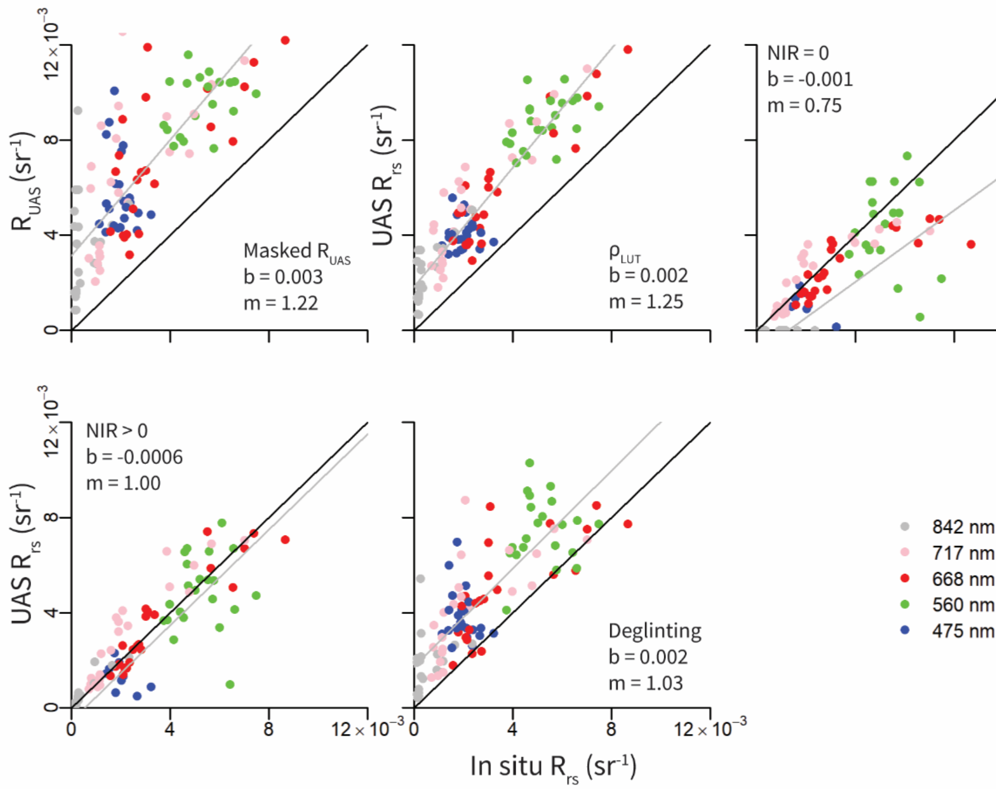


Figure 8. Comparison of UAS radiometry and *in situ* R_{rs} in all bands at each station (n=28) using different methods to remove surface reflected light after initial sun glint masking. A) Total UAS-derived reflectance (R_{UAS}), B) ρ look-up table from HydroLight simulations, C) Dark pixel assumption with NIR = 0, D) Dark pixel assumption with NIR > 0, E) Deglinting methods following Hochberg et al. (2003). Negative values are not shown in plots.

Water quality retrieval algorithms

R_{rs} measurements with L_{SR} estimated using a baseline NIR value (NIR>0) performed best when compared to *in situ* chlorophyll *a* data (R^2 = 0.37, RMSE = 5.89, RRMSE = 37%, $p < 0.05$). A stepwise model selection by AIC demonstrated that the green, red edge, and NIR bands were most important in estimating chlorophyll *a* concentration (Figure 9, $R^2 = 0.42$, RMSE = 5.90, RRMSE = 37%, $p < 0.05$) and a remotely sensed chlorophyll *a* algorithm was determined as:

$$Chlorophyll\ a\ (\mu g/L) = 24.02 + R_{rs}(560) * -4337.88 + R_{rs}(717) * 9639.75 + R_{rs}(842) * -2922.80$$

R_{rs} measurements with the ‘deglinting’ technique performed best when compared to *in situ* TSS data, ($R^2 = 0.72$, RMSE = 2.53, RRMSE = 9%, $p < 0.05$). A stepwise model selection by AIC demonstrated that the blue, red, red edge, and NIR bands were most important in estimating TSS concentration (Figure 9, $R^2 = 0.73$, RMSE = 2.53, RRMSE = 9%, $p < 0.05$) and a remotely sensed TSS algorithm was determined as:

$$TSS\ (mg/L) = 30.57 + R_{rs}(475) * 1364.86 + R_{rs}(668) * -5255.88 + R_{rs}(717) * 2548.08 + R_{rs}(842) * 4579.36$$

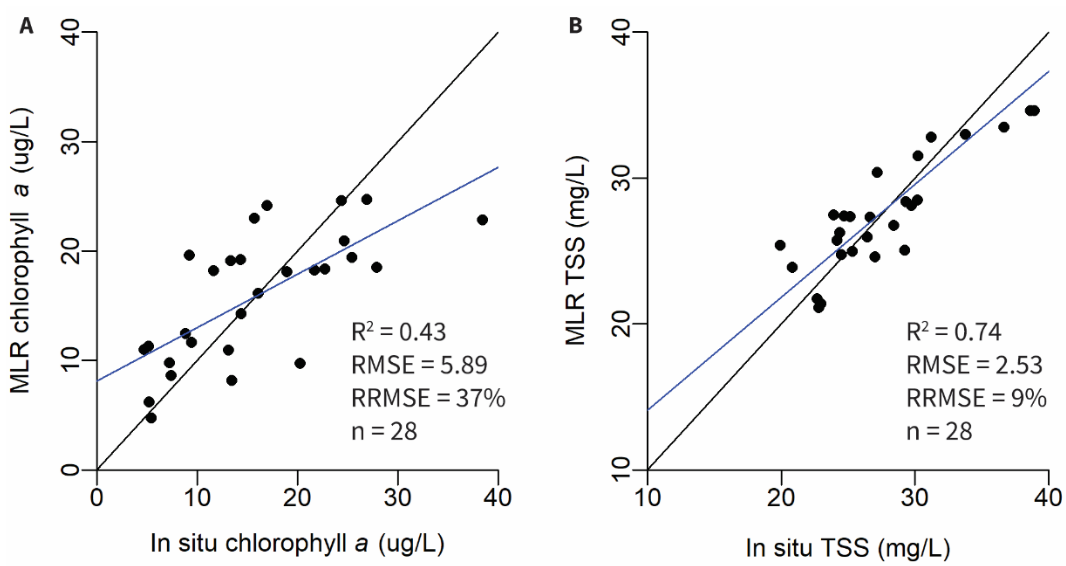


Figure 9. Comparison between *in situ* and modelled A) chlorophyll *a* concentration and B) TSS concentration from multiple linear regressions

Algorithms were applied to respective UAS-derived R_{rs} values and average chlorophyll *a* and TSS concentrations were mapped, along with mosaiced georeferenced TIFFs of individual image captures (Figure 10). Average chlorophyll *a* concentration tended to increase at upriver sites while average TSS concentrations were higher at downriver sites (Figure 10). This trend was also seen in the *in situ* data. Within the mosaic of individual TIFFs, slight variations in chlorophyll *a* and TSS concentration are visible (Figure 10).

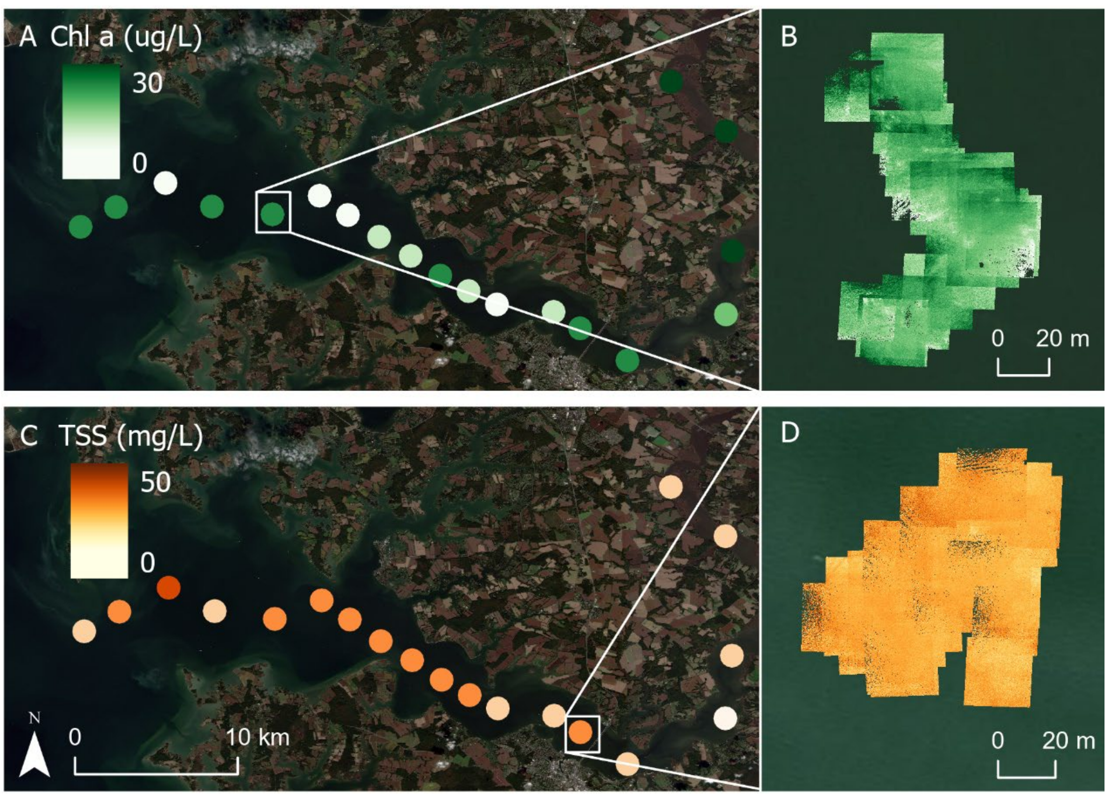


Figure 10. A) Average chlorophyll *a* concentration across image captures (n=30) at each station. B) Example of mosaiced georeferenced TIFFs collected at one station with chlorophyll *a* algorithm applied. C) Average TSS concentration across image captures (n=30) at each station. D) Example of mosaiced georeferenced TIFFs (n=30) collected at one station with TSS algorithm applied

Discussion

This study is one of the first to compare methods to remove surface reflected light from high resolution UAS multispectral measurements. An initial filtering procedure and four methods to remove L_{SR} were evaluated against *in situ* R_{rs} measurements. Water quality algorithm performance is compared to those found in the literature. Results have broad implications for improving UAS-derived water quality measurements in coastal waters, but several aspects and caveats of this study merit additional discussion.

Performance of methods to remove surface reflected light (L_{SR})

Removing surface reflected light from the total radiance measured by a UAS ensures only water-leaving information is used as input into water quality retrievals. Initial filtering techniques successfully masked instances of specular sun glint and non-water objects (i.e., boat) in the UAS imagery using empirical upper and lower R_{rs} limits. This primary step ultimately improved final R_{rs} measurements and derived water quality products. While the magnitude of L_{sky} can impact L_{SR}, the spectral dependence ($\sim\lambda^{-4}$) is constrained and can be measured; therefore, the effective sea-surface reflectance of a wave facet (ρ) needs to be accounted for. Nonetheless, it is important to consider the impact of the varying absolute magnitude of L_{sky} and future work should analyze the effect of L_{sky} variability on UAS R_{rs} measurements.

The ρ_{LUT} method utilized a ρ lookup table approach developed from HydroLight simulations where ρ values were obtained depending on the solar zenith angle, wind speed, and cloud cover at the time of UAS data collection. Mobley's (1999) HydroLight simulations led to recommendations of specific sensor viewing angles with a corresponding constant ρ value to reduce effects of L_{SR}; however, most consumer grade UAS sensors and mounts do not have the ability to change viewing angles and are fixed at a nadir viewing angle. Therefore, the ρ values in the look-up table correspond to nadir viewing angles. This method resulted in UAS R_{rs} values that were generally greater than *in situ* R_{rs} values, indicating L_{SR} was underestimated. This can most likely be attributed to a constant ρ value that was used to estimate L_{SR} across all pixels in each image, which averaged out the multifaceted characteristic of a water surface.

The NIR=0 and NIR>0 methods took advantage of the NIR waveband on the multispectral sensor and incorporated aspects of the dark pixel assumption which allowed L_{SR} to be removed from visible wavelengths by assuming that 1) a NIR signal is composed of L_{SR} and/or a spatially homogeneous NIR component in the water and 2) the amount of L_{SR} in the NIR band is linearly related to the amount of L_{SR} in the visible bands (Mobley, 1994; Siegel et al., 2000). Unlike the ρLUT method that applies a single value to an entire image, the dark pixel assumption approaches estimate L_{SR} on a per pixel basis which decreases resultant R_{rs} variability (Figure 6, 8). When NIR was assumed to be entirely negligible, calculated R_{rs} values were generally lower compared to *in situ* R_{rs} which can be attributed to an enhancement of NIR due to the presence of scattering inorganic particles in turbid waters. Thus, when a baseline NIR value was estimated, the method performed better (higher R^2, lower RMSE and RRMSE); however, negative R_{rs} values in the lower wavelengths indicated L_{SR} was still likely overestimated at some stations.

The 'deglinting' method, modelled off a glint removal algorithm introduced by Hochberg et al. (2003), and later Hedley et al. (2005), calculated a constant 'ambient' NIR brightness level which was removed from all pixels in all wavelengths. This approach resulted in UAS R_{rs} measurements slightly overestimated compared to *in situ* values but produced the highest R^2 and lowest RMSE of all methods. This method was applied by averaging pixel values across all thirty images collected at each station and is likely to perform better when applied on individual images.

The ρ_{LUT} method uses ρ values that are based on probability distributions of sea surface slopes which are related to the scale of satellite pixel resolution (100-1000m) (Cox & Munk, 1954). For images with much higher pixel resolution (>1-10m), statistical assumptions about the surface of water composed of many reflecting facets are less likely to hold (Kay et al., 2009). Therefore, to collect accurate water-leaving reflectance measurements from high resolution UAS imagery, it is recommended to use a pixel-based approach exploiting the high absorption of water at NIR wavelengths to estimate and remove L_{SR}. If a baseline NIR measurement can be retrieved, the dark pixel assumption should be used to remove L_{SR}. Otherwise, the 'deglinting' methods following Hochberg et al. (2003) and Hedley et al. (2005) are recommended. In open ocean waters without a strong influence of optically active properties, a pixel-based approach assuming NIR is negligible is expected to perform well.

It is important to note a potential caveat to the L_{SR} method evaluation. In this study, *in situ* R_{rs} measurements were provided by hyperspectral radiometers with a skylight blocking approach (Ahn, 1999; Lee et al., 2019). This approach consists of attaching an open-ended apparatus, or tube, to the front of a downward looking radiance radiometer and lowering it a few centimeters into the water, blocking surface-reflected light and allowing for a direct measurement of L_W. It is important to note that measurements from this technique are subject to instrument self-shading, which is a function of the water's optical properties, sun elevation, and the size of the skylight-blocking cone (Zhang et al., 2017; Lee et al., 2019). Zhang et al. (2017) estimated that self-shading accounts for approximately 1% to 20% error under most water properties and solar positions. Methods to correct for this self-shading have been derived (Zhang et al., 2017; Yu et al., 2021) and if applied, have the potential to improve relationships with UAS R_{rs} measurements.

Performance of water quality algorithms

Performance of the multiple linear regressions developed in this study were compared to existing chlorophyll *a* and TSS algorithms designed for coastal waters to determine if UAS measurements can produce accuracy within the range of other water quality algorithms. Performance of the UAS-derived chlorophyll *a* multiple linear regression (R^2 = 0.43, RRMSE = 37%) is comparable to other chlorophyll *a* algorithms found in the literature (Gitelson et al., 2008; Ruddick et al., 2001; Gons et al., 2002). A three-band chlorophyll *a* algorithm calibrated using a variety of coastal waters, including the Choptank River, resulted in a RRMSE of 51.9% (Gitelson et al., 2008), a two-band algorithm (red/NIR) with adaptive optimization in the second band calibrated with measurements from the North Sea and Lake Ijless, Netherlands resulted in a RRMSE of 37% (Ruddick et al., 2001), and a two-band algorithm (red/NIR) designed for the Medium Resolution Imaging Spectrometer (MERIS) satellite sensor and calibrated using a variety of coastal and inland waters including Lake IJssel (Netherlands), the Chinese Lake Tau Hu, and the Hudson/Raritan Estuary (New York/New Jersey) resulted in a standard error of 9.2 μg/L (Gons et al., 2002). Performance of the UAS-derived TSS multiple linear regression (R^2 = 0.73, RMSE = 2.53, RRMSE = 9%) is also comparable to existing TSS algorithms found in the literature (Novoa et al., 2017; Nechad et al., 2010). Algorithms including a single-band (NIR) second-order polynomial and single-band (red/green) linear models calibrated with measurements from the Gironde Estuary, France resulted in RRMSE values ranging from of 9.11% to 16.41% (Novoa et al., 2017) and a non-linear regression calibrated with measurements from the Southern North Sea resulted in

RRMSE values less than 30% (Neched et al., 2010). Future work will include improving water quality algorithms.

UAS sensor considerations

The innovative use of UAS technology for environmental research is a relatively new field and researchers are only beginning to understand and alleviate the various methodological and sensor performance challenges. Sensors degrade over time from use and environmental conditions which can impact the accuracy of the data being collected. The most recent MicaSense model, the MicaSense RedEdge-MX, is integrated with low-cost, image-frame complementary metal-oxide semiconductor (CMOS) sensors, which compared to typical charged-coupled device (CCD) sensors, tend to generate more noise and have lower sensitivity levels (Mamaghani & Savaggio, 2019). In the present study, raw values were radiometrically calibrated using a workflow[7] provided by Micasense which implements default metadata parameters that remain the same unless a new factory calibration is performed. Due to sensor degradation, these values are likely to gradually decline over time, lessening the accuracy of the radiometric calibration processing (Mamaghani & Savaggio, 2019). Studies have improved sensor performance by performing vicarious radiometric calibration using ground targets and panels with known radiometric accuracy, calibrating sensors using National Institute of Standards (NIST)-traceable equipment in a laboratory, and developing look-up tables for correction factors to update calibration parameters (Del Pozo et al., 2014;

Mamaghani & Salvaggio, 2019; Cao et al., 2020). Baek et al. (2021) conducted an assessment on radiometric accuracy for the MicaSense RedEdge-MX sensor by comparing data to hyperspectral sensors with NIST-traceable calibration (TriOS RAMSES) and showed that MicaSense RedEdge-MX radiance is approximately 5 to 16% lower, and irradiance is approximately 1 to 20% lower, depending on wavelength (Baek et al., 2021). The radiometric accuracy of a new or recently/vicariously calibrated UAS sensor should meet the required radiometric accuracy of 5% that is expected with ocean color satellites (McClain et al., 1992).

Precise registration of multispectral bands within a UAS image capture is also important to derive accurate spectral radiometric values across pixels. Sources of misregistration include a difference in the lens location for each band, image acquisition times, and exposure times which can all influence R_{rs} and resulting water quality variables (Kim et al., 2020). In the present study, images collected in each band were registered using MicaSense's default alignment function[7]. This three step process unwarps images using built-in lens calibration, determines a transformation to align each band to a common band, and crops pixels which do not overlap in all bands. Although this method seemed to perform well in this study; it is acknowledged that this type of band registration can perform poorly with images of locational errors, such as moving water, and can produce noise even after removing surface reflected light (Kim et al., 2020). Kim et al. (2020) developed a novel morphological band registration technique designed for high resolution water quality analysis which effectively removes misregistration noise and improves the accuracy of R_{rs}. Future aquatic UAS remote sensing work should consider adapting this technique to improve UAS remotely sensed retrievals over water.

Caveats and considerations

Many UAS aquatic remote sensing studies use Structure-from-Motion (SfM) photogrammetric techniques to stitch individual UAS images into ortho- and georectified mosaics (Arango et al., 2019; Olivetti et al., 2020; McEliece et al., 2020; Castro et al., 2020). This approach applies matching key points from overlapping UAS imagery in camera pose estimation algorithms to resolve 3D camera location and scene geometry (Westoby et al., 2012; Arango et al., 2020). Commonly used software (e.g., Pix4D) provide workflows that radiometrically calibrate, georeference, and stitch individual UAS images using a weighted average approach to create at-sensor reflectance 2D orthomosaics (Olivetti et al., 2020). L_{SR} removal methods and water quality algorithms can be directly applied to reflectance orthomosaics to effectively derive water quality products of an entire water body. However, current photogrammetry techniques are not capable of stitching UAS images captured over large bodies of water due to a lack of key points in images of homogenous water surfaces (Arango et al., 2020). Orthomosaics of smaller water bodies or rivers can be created if UAS images contain enough surrounding land features containing keypoints that the photogrammetry software can use to successfully stitch the images containing water. This can be accomplished by increasing flight altitude, with the trade-off of lower spatial resolution. Alternative methods include a statistical interpolation method; however interpolated reflectance values can be imprecise when compared to true reflectance values (Arango et al., 2020).

Management implications and future research

Water quality monitoring is important for tracking water quality trends, identifying and mitigating pollution sources, and discerning potential human health risks. Traditional *in situ* based methods of sampling at discrete stations can be expensive due to high costs for boat time and analysis and can also potentially omit important water quality phenomena. Traditional satellite remote sensing can capture variability throughout time and space; however, limitations including the presence of clouds, atmospheric effects, land adjacency effects, and spatial resolution can hinder periodic monitoring (Shi & Wang, 2009; Becker et al., 2019). UAS fill an operational gap between *in situ* and satellite remote sensing methods. While the current available commercial multispectral UAS sensor technology is geared towards terrestrial applications, mostly precision agriculture, the spectral bands have been useful in retrieving water quality parameters in aquatic water bodies (Choo et al., 2018; Baek et al., 2019; Arango et al., 2019; Castro et al., 2020; Olivetti et al., 2020). A UAS sensor package designed for aquatic environments would undoubtedly improve remotely sensed retrievals and water quality measurements.

Since UAS are deployed at a low altitude, atmospheric corrections and remedies to the land adjacency effect are eliminated. UAS are rapidly deployable and can provide the spatial and temporal variability required for useful water quality monitoring in a dynamic and rapidly evolving environment. UAS can enhance fine-scale physical oceanography research by resolving small-scale phenomena and physical processes such as patchy algal blooms, frontal structures, and turbulence characteristics (Figure 11, Shang et al., 2017; Osadchiev et al., 2020). UAS remotely-sensed water quality retrievals will also likely improve with the development of lightweight, off-the-shelf hyperspectral sensors, allowing for higher spatial and spectral resolution to better distinguish optical properties of the water (Shang et al., 2017; O'Shea et al., 2020).

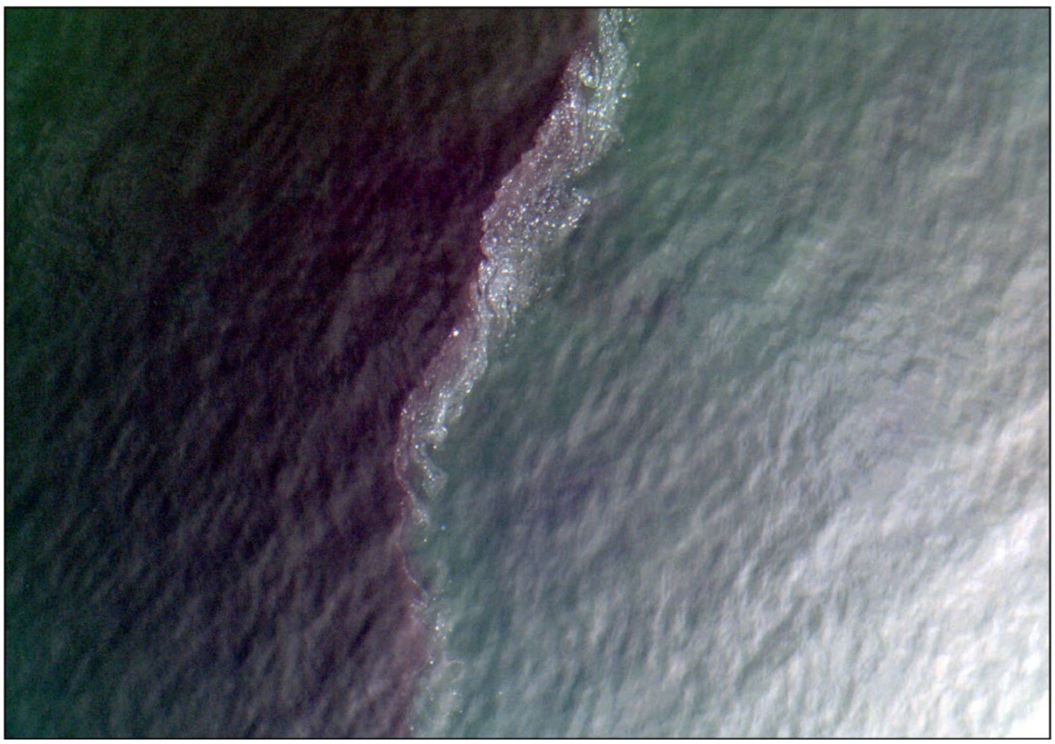


Figure 11. Frontal process observed in UAS RGB imagery collected in the Choptank River, Maryland, United States with MicaSense RedEdge-MX multispectral camera.

Future research in turbid coastal waters may benefit from sensor packages that use longer wavelengths (e.g., shortwave infrared, SWIR). SWIR wavelengths have shown to be more effective in satellite atmospheric correction techniques in turbid waters due to the stronger water absorption relative to NIR (Shi & Wang, 2009). This will be advantageous for removing L_{SR} and will improve R_{rs} retrievals in coastal, turbid waters. Future research could also consider combining UAS radiometry with photogrammetry computer vision to estimate L_{SR}. Scheider-Zapp et al. (2019) developed a method to estimate a hemispherical-directional reflectance factor (HDRF) from multi-angular UAS measurements. From a combination of photogrammetry and radiometry, a precise estimation of the downwelling light sensor position and orientation can be used to derive a multi-angular reflectance factor, which has the potential to significantly improve estimates of the multifaceted aspect of L_{SR}. An alternative approach to reduce L_{SR} is installing additional hardware to a UAS sensor. O'Shea et al. (2020) tested a hardware-based vertical polarizer on a hyperspectral spectrometer which effectively blocked horizontally polarized reflected skylight. This relatively simple approach could be an attractive solution to researchers and managers who are interested in applying a single algorithm to R_{rs} values to derive water quality parameters and should be further investigated.

Conclusion

UAS-based applications of multispectral or hyperspectral remote sensing in aquatic remote sensing have the potential to effectively fill current observation gaps in aquatic remote sensing and provide critical information needed for water quality forecasting, ecosystem monitoring, and ultimately climate change research. While atmospheric effects can usually be ignored in low altitude UAS flights, the effect of sun glint and surface reflected light should be accounted for in order to obtain the highest accuracy of water quality data. The inclusion of surface reflected light can lead to an overestimation of R_{rs} and remotely sensed water quality retrievals. This study presents a comparison of four approaches to remove sun glint and surface reflected light that can be applied to UAS remote sensing to derive water quality parameters such as chlorophyll-*a* and TSS concentration. Overall, the performance of the MicaSense RedEdge-MX multispectral sensor appears sufficient for providing high resolution water quality estimates of coastal water bodies when surface reflected light is removed. Of the four approaches examined, a pixel-based deglinting procedure utilizing the brightness of the NIR band performed best when compared to *in situ* R_{rs} measurements. This method also led to the best estimates of TSS while a pixel-based approach utilizing an ambient NIR signal to estimate and remove surface reflected light led to the best estimates of chlorophyll *a* concentration. Future work will include improving algorithms for water quality parameters. Future research should also consider the effects of sensor calibration and the residual misregistration between bands of a UAS multispectral camera.

References

Ahn, Y. H., Ryu, J. H., & Moon, J. E. (1999). Development of redtide & water turbidity algorithms using ocean color satellite. *KORDI Report No. BSPE*, 98721-00.

American Public Health Association (APHA). 1995. Standard Methods for the Examination of Water and Wastewater. 19th Ed. Topic 2540 Solids. APHA, Washington D.C.

Anderson, K., & Gaston, K. J. (2013). Lightweight unmanned aerial vehicles will revolutionize spatial ecology. *Frontiers in Ecology and the Environment*, *11*(3), 138-146.

Arar, E. J. And G. B. Collins. Method 445.0 In Vitro Determination of Chlorophyll a and Pheophytin ain Marine and Freshwater Algae by Fluorescence. U.S. Environmental Protection Agency, Washington, DC, 1997.

Arango, J. G., & Nairn, R. W. (2019). Prediction of Optical and Non-Optical Water Quality Parameters in Oligotrophic and Eutrophic Aquatic Systems Using a Small Unmanned Aerial System. *Drones*, *4*(1), 1.

Arango, J. G., Holzbauer-Schweitzer, B. K., Nairn, R. W., & Knox, R. C. (2020). Generation of geolocated and radiometrically corrected true reflectance surfaces in the visible portion of the electromagnetic spectrum over large bodies of water using images from a sUAS. *Journal of Unmanned Vehicle Systems*, *8*(3), 172-185.

Baek, J. Y., Jo, Y. H., Kim, W., Lee, J. S., Jung, D., Kim, D. W., & Nam, J. (2019). A New Algorithm to Estimate Chlorophyll-A Concentrations in Turbid Yellow Sea Water Using a Multispectral Sensor in a Low-Altitude Remote Sensing System. *Remote Sensing*, *11*(19), 2257.

Baek, S., Koh, S., & Kim, W. (2020). Calculation of correction coefficients for the RedEdge-MX multispectral camera through intercalibration with a hyperspectral sensor. *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography*, *38*(6), 707-716.

Becker, R. H., Sayers, M., Dehm, D., Shuchman, R., Quintero, K., Bosse, K., & Sawtell, R. (2019). Unmanned aerial system based spectroradiometer for monitoring harmful algal blooms: A new paradigm in water quality monitoring. *Journal of Great Lakes Research*, *45*(3), 444-453.

Cao, H., Gu, X., Wei, X., Yu, T., & Zhang, H. (2020). Lookup Table Approach for Radiometric Calibration of Miniaturized Multispectral Camera Mounted on an Unmanned Aerial Vehicle. *Remote Sensing*, *12*(24), 4012.

Choo, Y., Kang, G., Kim, D., & Lee, S. (2018). A study on the evaluation of water-bloom using image processing. *Environmental Science and Pollution Research*, *25*(36), 36775-36780.

Cillero Castro, C., Domínguez Gómez, J. A., Delgado Martín, J., Hinojo Sánchez, B. A., Cereijo Arango, J. L., Cheda Tuya, F. A., & Díaz-Varela, R. (2020). An UAV and Satellite Multispectral Data Approach to Monitor Water Quality in Small Reservoirs. *Remote Sensing*, *12*(9), 1514.

Cox, C., & Munk, W. (1954). Measurement of the roughness of the sea surface from photographs of the sun's glitter. *Journal of the Optical Society of America*, *44*(11), 838-850.

Del Pozo, S., Rodríguez-Gonzálvez, P., Hernández-López, D., & Felipe-García, B. (2014). Vicarious radiometric calibration of a multispectral camera on board an unmanned aerial system. *Remote Sensing*, *6*(3), 1918-1937.

Dugdale, S. J., Kelleher, C. A., Malcolm, I. A., Caldwell, S., & Hannah, D. M. (2019). Assessing the potential of drone-based thermal infrared imagery for quantifying river temperature heterogeneity. *Hydrological Processes*, *33*(7), 1152-1163.

Fisher, T. R., Fox, R. J., Gustafson, A. B., Koontz, E., Lepori-Bui, M., & Lewis, J. (2021). Localized Water Quality Improvement in the Choptank Estuary, a Tributary of Chesapeake Bay. *Estuaries and Coasts*, 1-20.

Gerum, R. C., Richter, S., Winterl, A., Mark, C., Fabry, B., Le Bohec, C., & Zitterbart, D. P. (2019). CameraTransform: A Python package for perspective corrections and image mapping. *SoftwareX*, *10*, 100333.

Gitelson, A. A., Schalles, J. F., & Hladik, C. M. (2007). Remote chlorophyll-a retrieval in turbid, productive estuaries: Chesapeake Bay case study. *Remote Sensing of Environment*, *109*(4), 464-472.

Gray, P. C., Ridge, J. T., Poulin, S. K., Seymour, A. C., Schwantes, A. M., Swenson, J. J., & Johnston, D. W. (2018). Integrating drone imagery into high resolution satellite remote sensing assessments of estuarine environments. *Remote Sensing*, *10*(8), 1257.

Gordon, H. R., & Clark, D. K. (1980). Atmospheric effects in the remote sensing of phytoplankton pigments. *Boundary-Layer Meteorology*, *18*(3), 299-313.

Hedley, J. D., Harborne, A. R., & Mumby, P. J. (2005). Simple and robust removal of sun glint for mapping shallow-water benthos. *International Journal of Remote Sensing*, *26*(10), 2107-2112.

Hochberg, E. J., Andréfouët, S., & Tyler, M. R. (2003). Sea surface correction of high spatial resolution Ikonos images to improve bottom mapping in near-shore environments. *IEEE transactions on geoscience and remote sensing*, *41*(7), 1724-1729.

Hooker, S. B., Lazin, G., Zibordi, G., & McLean, S. (2002). An evaluation of above- and in-water methods for determining water-leaving radiances. *Journal of Atmospheric and Oceanic Technology*, *19*(4), 486-515.

Gitelson, A. A., Schalles, J. F., & Hladik, C. M. (2007). Remote chlorophyll-a retrieval in turbid, productive estuaries: Chesapeake Bay case study. *Remote Sensing of Environment*, *109*(4), 464-472.

Gons, H. J., Rijkeboer, M., & Ruddick, K. G. (2002). A chlorophyll-retrieval algorithm for satellite imagery (Medium Resolution Imaging Spectrometer) of inland and coastal waters. *Journal of Plankton Research*, *24*(9), 947-951.

Gordon H.R., Clark D.K. (1980). Atmospheric effects in the remote sensing of phytoplankton pigments. Boundary-layer Meteorology 18: 299-313.

Gordon, H. R., & Wang, M. (1994). Retrieval of water-leaving radiance and aerosol optical thickness over the oceans with SeaWiFS: a preliminary algorithm. *Applied optics*, *33*(3), 443-452.

Johnston, D. W. (2019). Unoccupied aircraft systems in marine science and conservation. *Annual review of marine science*, *11*, 439-463.

Kay, S., Hedley, J. D., & Lavender, S. (2009). Sun glint correction of high and low spatial resolution images of aquatic scenes: a review of methods for visible and near-infrared wavelengths. *Remote sensing*, *1*(4), 697-730.

Kim, W., Jung, S., Moon, Y., & Mangum, S. C. (2020). Morphological band registration of multispectral cameras for water quality analysis with unmanned aerial vehicle. *Remote Sensing*, *12*(12), 2024.

Lavender, S. J., Pinkerton, M. H., Moore, G. F., Aiken, J., & Blondeau-Patissier, D. (2005). Modification to the atmospheric correction of SeaWiFS ocean colour images over turbid waters. *Continental Shelf Research*, *25*(4), 539-555.

Lee, Z., Ahn, Y. H., Mobley, C., & Arnone, R. (2010). Removal of surface-reflected light for the measurement of remote-sensing reflectance from an above-surface platform. *Optics Express*, *18*(25), 26313-26324.

Lee, E., Yoon, H., Hyun, S. P., Burnett, W. C., Koh, D. C., Ha, K., ... & Kang, K. M. (2016). Unmanned aerial vehicles (UAVs)-based thermal infrared (TIR) mapping, a novel approach to assess groundwater discharge into the coastal zone. *Limnology and Oceanography: Methods*, *14*(11), 725-735.

Lee, Z., Wei, J., Shang, Z., Garcia, R., Dierssen, H. M., Ishizaka, J., & Castagna, A. (2019). On-water radiometry measurements: skylight-blocked approach and data processing. *Appendix to Protocols for Satellite Ocean Colour Data Validation: In Situ Optical Radiometry. IOCCG Ocean Optics and Biogeochemistry Protocols for Satellite Ocean Colour Sensor Validation, Volume 3.0*, 7.

Mamaghani, B., & Salvaggio, C. (2019). Multispectral sensor calibration and characterization for sUAS remote sensing. *Sensors*, *19*(20), 4453.

McEliece, R., Hinz, S., Guarini, J. M., & Coston-Guarini, J. (2020). Evaluation of Nearshore and Offshore Water Quality Assessment Using UAV Multispectral Imagery. *Remote Sensing*, *12*(14), 2258.

McClain, C. et al. (1992). SeaWiFS calibration and validation plan. In: NASA Technical Memorandum 104566, ed. by S. Hooker and E. Firestone. Vol. 3. Prelaunch Technical Report Series. NASA Goddard Space Flight Center, Greenbelt, Maryland, 41 pp

Mobley, C. D. (1999). Estimation of the remote-sensing reflectance from above-surface measurements. *Applied optics*, *38*(36), 7442-7455.

Mobley, C. D. (2015). Polarized reflectance and transmittance properties of windblown sea surfaces. *Applied optics*, *54*(15), 4828-4849.

Morgan, B. J., Stocker, M. D., Valdes-Abellan, J., Kim, M. S., & Pachepsky, Y. (2020). Drone-based imaging to assess the microbial water quality in an

irrigation pond: A pilot study. *Science of The Total Environment*, *716*, 135757.
Nechad, B., Ruddick, K. G., & Park, Y. (2010). Calibration and validation of a generic multisensor algorithm for mapping of total suspended matter in turbid waters. *Remote Sensing of Environment*, *114*(4), 854-866.
Novoa, S., Doxaran, D., Ody, A., Vanhellemont, Q., Lafon, V., Lubac, B., & Gernez, P. (2017). Atmospheric corrections and multi-conditional algorithm for multi-sensor remote sensing of suspended particulate matter in low-to-high turbidity levels coastal waters. *Remote Sensing*, *9*(1), 61.
Olivetti, D., Roig, H., Martinez, J. M., Borges, H., Ferreira, A., Casari, R., ... & Malta, E. (2020). Low-Cost Unmanned Aerial Multispectral Imagery for Siltation Monitoring in Reservoirs. *Remote Sensing*, *12*(11), 1855.
Osadchiev, A., Barymova, A., Sedakov, R., Zhiba, R., & Dbar, R. (2020). Spatial structure, short-temporal variability, and dynamical features of small river plumes as observed by aerial drones: case study of the Kodor and Bzyp river plumes. *Remote Sensing*, *12*(18), 3079.
O'Shea, R. E., Laney, S. R., & Lee, Z. (2020). Evaluation of glint correction approaches for fine-scale ocean color measurements by lightweight hyperspectral imaging spectrometers. *Applied optics*, *59*(7), B18-B34.
Ruddick, K. G., Gons, H. J., Rijkeboer, M., & Tilstone, G. (2001). Optical remote sensing of chlorophyll a in case 2 waters by use of an adaptive two-band algorithm with optimal error properties. *Applied optics*, *40*(21), 3575-3585.
Ruddick, K. G., De Cauwer, V., Park, Y. J., & Moore, G. (2006). Seaborne measurements of near infrared water-leaving reflectance: The similarity spectrum for turbid waters. *Limnology and Oceanography*, *51*(2), 1167-1179.
Schneider-Zapp, K., Cubero-Castan, M., Shi, D., & Strecha, C. (2019). A new method to determine multi-angular reflectance factor from lightweight multispectral cameras with sky sensor in a target-less workflow applicable to UAV. *Remote Sensing of Environment*, *229*, 60-68.
Shang, S., Lee, Z., Lin, G., Hu, C., Shi, L., Zhang, Y., ... & Yan, J. (2017). Sensing an intense phytoplankton bloom in the western Taiwan Strait from radiometric measurements on a UAV. *Remote Sensing of Environment*, *198*, 85-94
Shang, Z., Lee, Z., Dong, Q., & Wei, J. (2017). Self-shading associated with a skylight-blocked approach system for the measurement of water-leaving radiance and its correction. *Applied optics*, *56*(25), 7033-7040.
Shi, W., & Wang, M. (2009). An assessment of the black ocean pixel assumption for MODIS SWIR bands. *Remote Sensing of Environment*, *113*(8), 1587-1597.
Siegel, D. A., Wang, M., Maritorena, S., & Robinson, W. (2000). Atmospheric correction of satellite ocean color imagery: the black pixel assumption. *Applied optics*, *39*(21), 3582-3591.
Spyrakos, E., O'Donnell, R., Hunter, P. D., Miller, C., Scott, M., Simis, S. G., ... & Tyler, A. N. (2018). Optical types of inland and coastal waters. *Limnology and Oceanography*, *63*(2), 846-870.

Su, T. C., & Chou, H. T. (2015). Application of multispectral sensors carried on unmanned aerial vehicle (UAV) to trophic state mapping of small reservoirs: a case study of Tain-Pu reservoir in Kinmen, Taiwan. *Remote Sensing*, *7*(8), 10078-10097.

Su, T. C. (2017). A study of a matching pixel by pixel (MPP) algorithm to establish an empirical model of water quality mapping, as based on unmanned aerial vehicle (UAV) images. *International journal of applied earth observation and geoinformation*, *58*, 213-224.

Totsuka, S., Kageyama, Y., Ishikawa, M., Kobori, B., & Nagamoto, D. (2019). Noise removal method for unmanned aerial vehicle data to estimate water quality of Miharu dam reservoir, Japan. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, *23*(1), 34-41.

Tzortziou, M., Herman, J. R., Gallegos, C. L., Neale, P. J., Subramaniam, A., Harding Jr, L. W., & Ahmad, Z. (2006). Bio-optics of the Chesapeake Bay from measurements and radiative transfer closure. *Estuarine, Coastal and Shelf Science*, *68*(1-2), 348-362.

Westoby, M. J., Brasington, J., Glasser, N. F., Hambrey, M. J., & Reynolds, J. M. (2012). 'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology*, *179*, 300-314.

Werdell, P. J., McClain, C. R. (2019). Satellite Remote Sensing: Ocean Color. *Encyclopedia of Ocean Sciences,* 3, 443-455.

Windle, A. E., Poulin, S. K., Johnston, D. W., & Ridge, J. T. (2019). Rapid and accurate monitoring of intertidal oyster reef habitat using unoccupied aircraft systems and structure from motion. *Remote Sensing*, *11*(20), 2394.

Xu, F., Gao, Z., Jiang, X., Shang, W., Ning, J., Song, D., & Ai, J. (2018). A UAV and S2A data-based estimation of the initial biomass of green algae in the South Yellow Sea. *Marine pollution bulletin*, *128*, 408-41

Yu, X., Lee, Z., Shang, Z., Lin, H., & Lin, G. (2021). A simple and robust shade correction scheme for remote sensing reflectance obtained by the skylight-blocked approach. *Optics Express*, *29*(1), 470-486.

Zeng, C., Richardson, M., & King, D. J. (2017). The impacts of environmental variables on water reflectance measured using a lightweight unmanned aerial vehicle (UAV)-based spectrometer system. *ISPRS Journal of Photogrammetry and Remote Sensing*, *130*, 217-230.

Chapter 5: Multi-temporal high-resolution marsh vegetation mapping using unoccupied aircraft system remote sensing and machine learning

Introduction

Coastal wetlands, existing at the nexus of land and water, are among the most productive ecosystems in the world and provide a suite of ecosystem services including the ability to sequester carbon, improve water quality, protect coastal regions from storms, and provide nursery grounds that support commercial fisheries (Barbier et al., 2011). Close to sea-level, coastal wetlands are vulnerable to anthropogenic and climate impacts; many of the world's wetlands have retreated due to land conversion and rising sea levels (Pendleton et al., 2012; Kirwan and Megonigal, 2013). However, important feedbacks between vegetation growth, water level and geomorphology allow wetlands to persist under rising sea levels (Feagin et al., 2010; Kirwan and Megonigal, 2013; Beckett et al., 2016; Kirwan et al., 2016; Alizad et al., 2018; Schieder et al., 2018; Flester and Blum, 2020). Long-term monitoring elucidates emerging vegetation and soil feedbacks, allowing marsh restoration and management practitioners to better mitigate threats of sea level rise and habitat loss.

Wetlands typically have a noticeable zonation of vegetation controlled by physiological constraints and competitive displacement due to physical stress, resource competition, and nutrient availability (Bertness, 1991; Aerts, 1999; Emery et al., 2001). For example, tidal flooding establishes a stress gradient with soil anoxia, salinity, and water inundation decreasing towards the terrestrial border of a marsh (Mendelssohn et al., 1981). Low marsh vegetation, such as the cordgrass *Spartina alterniflora* contain extensive aerenchyma for increased gas exchange to persist in waterlogged environments (Bertness, 1991) as well as deeper root profiles that provide higher stability in surging water (Howes et al., 2010). In times of persistent tidal inundation from higher rates of sea level rise, wetlands can adapt by vertically accreting sediment; however, if accretion rates are low due to a lack of sedimentation or carbon burial, marsh vegetation tends to shift and migrate inland, a process known as marsh migration or transgression (Feagin et al., 2010; Kirwan and Megonigal, 2013; Enwright et al., 2016; Kirwan et al., 2016; Alizad et al., 2018; Schieder et al., 2018). Globally, coastal wetlands are responding to sea level rise and increased inundation by migrating upslope, leading to substantial shifts in zonation, habitat loss, and pond expansion (Feagin et al., 2010; Qi et al., 2020). Long term monitoring of marsh vegetation can improve our understanding of how tidal wetlands respond to sea level rise and assist in restoration management.

For decades, spectral and structural characteristics from high resolution satellite and aircraft imagery have been used to monitor and map spatially and temporally complex wetlands (Gross et al., 1987; Klemas, 2013; Byrd et al., 2014; Evans et al., 2014; Lane et al., 2014; Massetti et al., 2016; Liu et al., 2017; Qi et al., 2020). Using field measurements for training and validation, vegetation classification models are used to monitor large-scale marsh habitat over time (Mahdavi et al., 2018). However, the spatial resolution of satellite and airborne imagery, albeit very high (~1–30 m), can still result in a large fraction of mixed pixels, leading to challenges in differentiating between vegetation types. Distinguishing between marsh vegetation species can also be challenging due to similar spectral signatures. Artigas and Yang (2005); Artigas and Yang (2006) compared *in situ* leaf-level hyperspectral reflectance of *S. alterniflora* and *Spartina patens* and found no significant differences in most of the visible and near-infrared (NIR) region. Some studies have demonstrated that incorporating an image texture metric (i.e., the spatial variance of image grayscale levels of one spectral band) can improve marsh vegetation classifications by incorporating characteristic patchiness and other structural features into the distinguishing features of each vegetation type (Dronova et al., 2012; Lane et al., 2014). Additional data layers such as elevation data from Light Detection and Ranging (LiDAR) have also improved classification (Halls and Costin, 2016; Qi et al., 2020). Some LiDAR-derived DEMs, however, tend to overestimate the marsh platform elevation in marsh environments due to an inability to penetrate dense canopies (Hladik and Alber, 2012; Elmore et al., 2016). Additionally, synthetic aperture radar (SAR) imagery has also been successfully used for classifying and mapping marsh vegetation, proving to be a useful method in dense vegetation or with frequent cloud cover (Silva et al., 2010; Evans et al., 2014). However, the spatial resolution of SAR imagery can still lead to misclassifications of small-scale wetlands.

The emerging technology of unoccupied aircraft systems (UAS, or drones) combined with Structure from Motion (SfM) photogrammetry is a promising method for classifying and assessing complex wetland habitats with high accuracy (Wan et al., 2014; Kalacska et al., 2017; Doughty and Cavanaugh, 2019; Durgan et al., 2020; Ridge and Johnston, 2020; Nardin et al., 2021). UAS can provide ultra-high resolution (<5 cm) overlapping two- dimensional (2D) images that SfM photogrammetry software can then reconstruct into 2D orthomosaics and three-dimensional (3D) point clouds (Westoby et al., 2012). UAS-SfM photogrammetry can be a low-cost alternative to expensive high-resolution satellite or occupied aircraft approaches to obtain 2D and 3D radiometric measurements of an environment. The fusion of multispectral UAS imagery and SfM-derived digital surface models (DSMs) have been applied to monitor coastal morphology over time (Seymour et al., 2018), improve species-specific wetland habitat classifications (Gray et al., 2018), monitor marsh geomorphology (Kalacska et al., 2017), and provide an estimate of vegetation height (DiGiacomo et al., 2020).

Engineered ecosystems such as tidal marsh creation or restoration and living shoreline placement are a means to mitigate anthropogenic climate change by replacing or conserving lost habitat (Duarte et al., 2013). These climate mitigation tools are becoming increasingly prevalent but often without a comprehensive understanding of the ecosystem due to a lack of effective monitoring (Li et al., 2018). The on-demand capabilities of UAS can be adopted in restoration efforts to improve site assessment and provide long-term, real-time monitoring (Ridge and Johnston, 2020). The overarching objective of this study is to develop novel methods to improve species-specific marsh vegetation differentiation using high resolution UAS imagery and SfM techniques to track marsh vegetation shifts. Specifically, this study aims to 1) examine long term (17 years) and short term (4 years) changes in vegetation species distribution in a restored wetland, 2) determine the optimal time of year for high classification accuracy of the dominant species, *S. alterniflora* and *S. patens*, and 3) determine what variables are most important in marsh vegetation classifications. The results of this study demonstrate that interannual variability of marsh vegetation can be non-linear and long-term monitoring is needed for an accurate understanding of marsh dynamics and improved marsh restoration practices.

Methods

Study site

The Paul S. Sarbanes Ecosystem Restoration Project at Poplar Island, located in mid-Chesapeake Bay, Maryland is a large-scale habitat restoration site that receives dredged material from the navigation channels approaching Baltimore Harbor. When completed, 68 million cubic yards of dredged material will be used to restore approximately 694 ha of tidal wetland, upland, and open-water embayment habitat (USACE, 2020). Tidal marshes are constructed following dredged material placement in containment cells ranging in size from approximately 12–20 ha (Staver et al., 2020). Tidal exchange is established *via* a tidal inlet in the exterior dike of each cell. The site has a mean tidal range of 0.49 m (NOAA Tides and Currents, station ID 8571892) and surface salinity ranges from 8.89 to 16.13 ppt (1985–2021, Station CB4.1E, Chesapeake Bay Program). The present study focused on Cell 3D, a 13 ha marsh which was developed in 2005. The marsh was planted with nursery grown *S.*

alterniflora in the low marsh zone and *S. patens* in the high marsh zone, with a target of maintaining 85%–90% coverage (USACE, 2020; Figure 1).

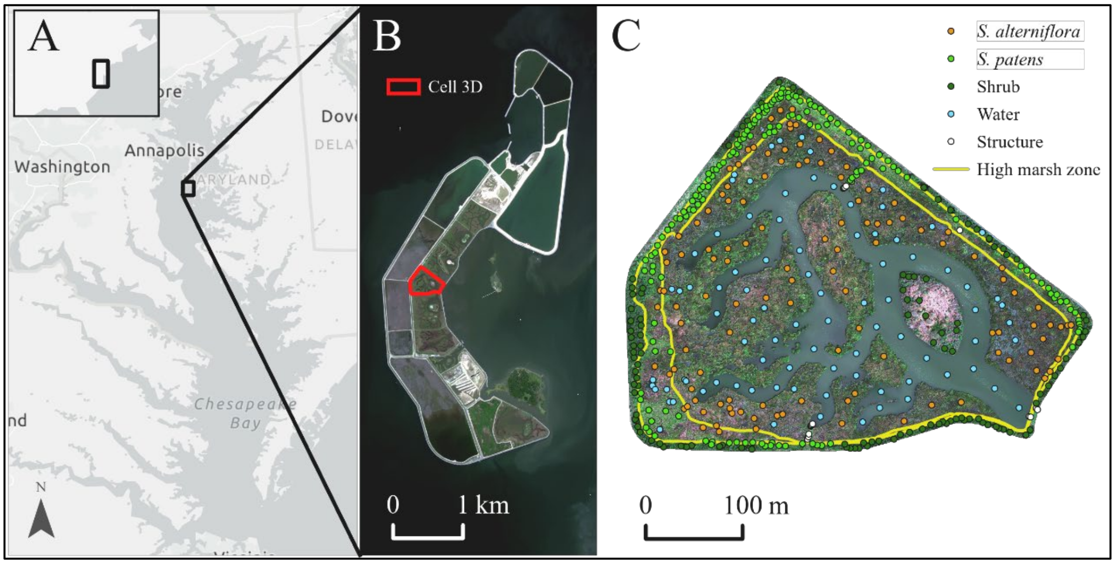


Figure 1. Overview of the study site location. (A) Location of Poplar Island in the Chesapeake Bay, Maryland, United States (B) WorldView-2 satellite image captured on 15 September 2020 of the Paul S. Sarbanes Ecosystem Restoration Project at Poplar Island with Cell 3D outlined in red (C) UAS RGB composite of Cell 3D collected on 26 August 2020 overlaid with the as-built planting boundaries of high and low marsh in yellow and field validation data used to train the random forest classification.

UAS description

Multispectral UAS data was collected with a MicaSense RedEdge-MX sensor (MicaSense, Seattle, Washington, United States). The sensor is a 8.7 × 5.9 × 4.5 cm multispectral camera capable of capturing five simultaneous bands on the electromagnetic spectrum in 12 bit radiometric resolution: blue (475 nm center, 32 nm full width half maximum, FWHM), green (560 nm center, 27 nm FWHM), red (668 nm center, 14 nm FWHM), red edge (717 nm center, 12 nm FWHM), and NIR

(842 nm center, 57 nm FWHM). The sensor was mounted on a DJI Phantom 4 Pro UAS using a 3D-printed mount that assured a nadir viewing angle while in flight. The UAS also contains a downwelling light sensor (DLS) which measures downwelling hemispherical irradiance in the same spectral wavebands during in-flight image captures and contains a calibrated magnetometer to provide heading and orientation information which, when combined, corrects for varying illumination conditions. The DLS was mounted above the UAS to eliminate shading and collected incident light centered at a 0° zenith angle.

UAS imagery and ground control point collection

UAS data was collected over Cell 3D on five different dates (21 November 2019, 26 August 2020, 16 April 2021, 22 June 2022, and 7 November 2022) between 10:00a.m. and 2:00p.m. EST, with tides ranging from 0.29 to 1.1 m, referenced to North American Vertical Datum established in 1988 (NAVD88). A ~5 cm/pixel ground sampling distance (GSD) was achieved by collecting imagery at an average 76 m altitude above sea level with 70% longitudinal and 80% latitudinal overlap using the flight planning application Pix4Dcapture (Pix4D, Prilly, Switzerland). Six ground control points (GCPs) were surveyed in the field with a real-time kinematic global positioning system (RTK-GPS, Reach RS+, Emlid, Hong Kong). Three of the GCPs were made from high-density polyurethane black and white checkerboard tiles (0.0929 m^2) and situated on PVC pipes of differing lengths distributed around the 13 ha survey area (Supplementary Figure S1). The other three GCPs were locations of permanent structures (e.g., corner of piers). Bundle block adjustment results were strong across all datasets, with mean reprojection errors of 0.11, 0.10, 0.12, 0.11, 0.13 pixels, respectively and relative mean of geolocational accuracy was under 1 m in the *X*,*Y* directions and under 1.5 m in the Z directions for each dataset.

UAS imagery was processed with Pix4D Mapper Pro SfM photogrammetry software v4.6.4. GCP measurements were incorporated into the image processing in the software program by partial automation to georectify all UAS surveys to achieve the greatest positional and vertical accuracy. Horizontal data were referenced to the World Geodetic Datum 1984 (WGS 1984) Universal Transverse Mercator (UTM) Zone 18N, and vertical data were referenced to NAVD88. Reflectance values were obtained following the workflow described in Assmann et al. (2018). In summary, the MicaSense sensor collected radiance values which were stored as arbitrary digital numbers (DNs). These values were used to obtain reflectance by combining sensor and illumination information, as follows:

*Reflectance = digital number/(radiometrically calibrated pixel value * illumination)*

Within the Pix4D radiometric calibration process, sensor metadata parameters were used to apply several corrections such as correcting for vignetting (pixels on the outside of images receive less light than those in the center), compensating for dark current noise (sensitivity even when no photons enter device), and correcting for spectral overlap, sensor sensitivity, and image brightness (ISO, aperture, exposure time). Illumination values from the DLS and values collected from the calibrated reflectance panel were applied to the radiometrically calibrated values to account for variation caused by differences in ambient light due to weather and sun. This process outputs unitless surface reflectance values, ranging from 0 to 1 (Assmann et al., 2018). Blue, green, red, red edge, NIR reflectance orthomosaics were produced as well as a normalized difference vegetation index (NDVI) orthomosaic, calculated as the difference between NIR and red reflectance divided by their sum. Orthorectified digital surface models (DSMs) were also created by applying an inverse weighted distance (IDW) interpolation from the 3D classified point cloud,

Field data collection

Field data were collected to use as a training dataset using a range of methods. Field vegetation identification was collected at the time of each UAS flight using a RTK GPS (6–10 points per dataset, 0.01 m average lateral root mean square error). Additional vegetation identification points were collected throughout the study period by wetland restoration partners and used as training data for each respective flight. U.S. Fish and Wildlife Service personnel collected sixteen vegetation identification data in 2019 and 2021 across four transects spanning high marsh and low marsh vegetation; U.S. Army Corps of Engineers personnel collected annual vegetation identification of one low and one high marsh quadrat; and an ongoing vegetation biomass monitoring study (Staver et al., 2020) provided annual vegetation identification data in six high marsh and six low marsh locations. Despite this extensive field collection effort, the size of the training data was increased by 90%–95% through visual inspection of the high resolution UAS imagery which was reinforced by known identification of the field data. In total, habitat from 35 to 155 points (pixels) were identified for each class: *S. alterniflora*, *S. patens*, structure (any human made material), shrub (vegetation mainly on the surrounding dike), and water (Figure 1). For each dataset, all field data points from the multiple sources were randomly split into a training set (70%) to train a classification model and a test set (30%) for assessing the prediction accuracy of the models.

Multispectral reflectance spectra

For each UAS dataset, reflectance values from all bands were extracted from the *S. alterniflora* and *S. patens* field training pixels to calculate an average multispectral spectrum of each vegetation species. These spectra were compared across seasons to determine differences in vegetation reflectance. Additionally, following methods of Artigas and Young (2006), the red-edge first derivative, or slope, of reflectance at 668 and 717 nm was measured to further separate the two vegetation species.

Texture analysis

Image texture was derived from gray-level co-occurrence matrices (GLCM) ('glcm' package in R v 4.2.1). GLCM textural features are based on statistics which describe how often one gray tone appears in a specified spatial relationship to another gray tone which can reflect changes in uniformity, variation, and similarity in different directions and intervals (Haralick et al., 1973; Wang et al., 2018). Specifically, in this study the variance was chosen as the texture metric which can be represented by this sum of squares equation:

$$Variance = \sum_{i=1}^{N}\sum_{j=1}^{N}(i - \mu)^2 p(i,j)$$

Where N is the total number of gray levels in the image (NIR reflectance is scaled to 32 levels), $p(i,j)$ is the (i,j)-th entry of the normalized GLCM, that is $p(i,j) = P(i,j)/\mathrm{I},_{i,j} P(i,j)$, where $P(i,j)$ is the (i,j)-th entry of the computed GLCM, and μ is the mean of the rows and column sums of the GLCM. A high variance value stems from increased variation of reflectance intensity in the pixel window and is associated with jagged and rough textures, such as differing vegetation heights of *S. alterniflora,* while a low variance value is associated with smooth or fine textures such as the more planophile leaf orientation of *S. patens*. Average texture metrics were calculated over all directions (e.g., using shifts of 0°, 45°, 90°, and 135°) in a 43 × 43 moving pixel window (2.15 × 2.15 m) using one high-contrast band to lessen redundancy (NIR). The window size was chosen to best represent the average size of vegetation, following the methods of Feng et al. (2015).

Classification model generation

Random forest decision tree models (RandomForest package, R v 4.2.1) were used for the marsh vegetation classifications. A random forest model is an ensemble of classification and regression trees (CARTs) fed by bootstrapped training data (Breiman, 2001; Belgiu and Dragut, 2016). It has been used in numerous applications of wetland classifications (Mutanga et al., 2012; Elmore et al., 2016; Pricope et al., 2022) and has been shown to be a robust machine learning classification algorithm that overcomes overfitting due to the large number of decision trees (Breiman, 2001; Feng et al., 2015; Belgiu and Dragut, 2016; Berhane et al., 2018). A random forest implementation was generated for each UAS dataset by training 1,000 decision trees (*ntree* = 1,000) with an optimum number of features (*mtry*) calculated using the '*tuneRF*' function. Input into the random forest consisted of reflectance from the five wavebands, NDVI, elevation from the DSM, and texture (variance calculated from GLCM).

Since training data was collected concurrently with each survey, each classification was independently generated and validated against each unique survey. Each random forest model was fitted to raster data using the '*predict*' function to produce a classified output for each dataset. To remove some noise, a majority filter was applied that replaced cells based on the majority of their eight contiguous neighboring cells in a 3 × 3 pixel window. To assess the benefit of including the structural components of elevation and texture, random forest classifications were applied to just the five multispectral bands + NDVI, the five multispectral bands + NDVI + texture, and the five multispectral bands + NDVI + elevation. To assess the benefit of the high resolution of the UAS imagery, random forest classifications were applied to resampled UAS five bands at the following resolutions: 1.2, 2.4, 4, and 10 m.

Variable importance was measured using the Mean Decrease in Accuracy (MDA) which represents how much removing each variable reduces the accuracy of the model, also known as permutation importance (Breiman, 2001). This is done by computing the prediction error on the out-of-bag (OOB) portion of the data before and after permutation, averaging the difference over all trees, and normalizing by the standard deviation of the differences. Classification performance was analyzed using three different metrics. Overall accuracy expresses the percentage of map area that has been correctly classified when compared to reference data, calculated by dividing the total number of correctly classified pixels by the total number of reference pixels (Story and Congalton, 1986). To assess how accuracy was distributed across individual categories, Producer's and User's accuracy were calculated. Producer's accuracy is a performance measure of how well reference pixels are classified in each class and is calculated by dividing the number of correctly classified reference pixels in each class by the number of reference pixels known to be of that class. User's accuracy is a performance measure of how well pixels classified represent the true class and is calculated by dividing the number of correctly classified pixels in each class by the total number of pixels that were classified in that class. In brief, Producer's accuracy is a measure of omission error where pixels that have not been correctly classified have been omitted from the correct class and User's accuracy is a measure of commission error where pixels from the classified image do not represent that class on the ground (Story and Congalton, 1986).

Calculating vegetation change

To study temporal vegetation change, the UAS classifications were compared to the 2005 as-built boundaries where only *S. patens* was planted upslope of the high marsh/low marsh boundary (high marsh zone) and *S. alterniflora* was planted downslope of that boundary (low marsh zone) (Figure 1). The UAS classified maps were converted to a polygon and vegetation area (m^2) was calculated in the high marsh zone for each vegetation species across each dataset.

Results

Spectral and structural observations of marsh vegetation from UAS data

Average multispectral reflectance spectra demonstrates that *S. patens* exhibits higher reflectance across all wavebands than *S. alterniflora*, particularly in the longer wavebands (717 nm, 842 nm, Figure 2). The separation in the longer wavelengths is more apparent in the November, August, and April datasets compared to the June dataset, likely due to differences in chlorophyll levels between the two vegetation species during these seasons (Figure 2). Similarly, there are subtle peaks in the green waveband (560 nm) for both species in the June and August datasets while the spectra appear flatter across wavebands for the November and April datasets likely due to increased pigmentation during the summer months.

Figure 3 depicts broad differences between canopy elevation, NDVI, and red edge slope between species and seasons. Across all seasons, canopy elevation derived from the DSM of *S. alterniflora* and *S. patens* ranged from 0.03 m to 2.02 m and 0.12

to 2.56 m, respectively. The average *S. alterniflora* canopy elevation was highest in June (1.05 m), lowest in August (0.5 m) and the average *S. patens* canopy elevation was highest in June (1.16 m) and lowest in November (0.79 m and 0.85 m for 2019 and 2022, respectively). *S. alterniflora* and *S. patens* NDVI values range from 0.08 to 0.91 and0.14 to 0.93, respectively with June containing the highest average NDVI for *S. alterniflora* (0.91) and August containing the highest NDVI for *S. patens* (0.93). April contains the lowest NDVI value for both species (0.08 and 0.14, respectively). Red edge slope is on average higher for *S. patens* across all datasets ranging from 0.007 to 0.14 with the lowest values in April. Average red edge slope was highest for both species in the August and June datasets (Figure 2).

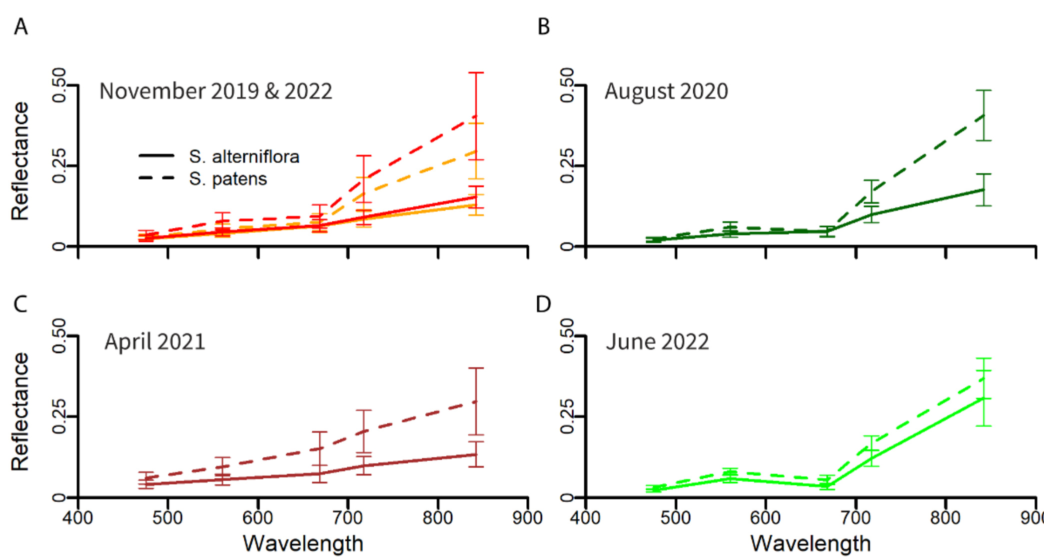


Figure 2. Multispectral reflectance spectra averaged from all *Spartina alterniflora* (solid line) and *Spartina patens* (dashed line) field training points across (A) November 2019 (orange) and November 2022 (red), (B) August 2020 (dark green), (C) April 2021 (brown), (D) June 2022 (light green). Error bars represent standard deviation of reflectance at each band.

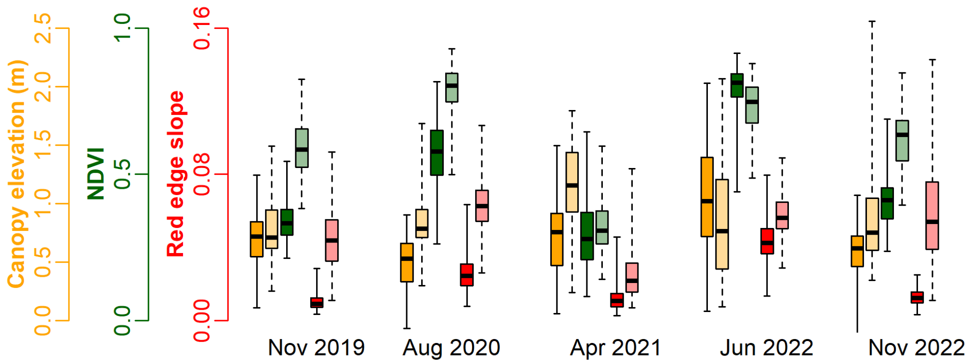


Figure 3. Boxplots of DSM canopy elevation (m) (yellow), NDVI (green), and the red edge slope (red) of *Spartina alterniflora* (dark hue) and *Spartina patens* (light hue) for each dataset.

Random forest classification

When applying all five UAS spectral bands, NDVI, texture, and elevation to each random forest classification, overall classification accuracy ranged from 97% to 99% across all datasets, with a mean of 98% (Supplementary Table S1). *S. alterniflora* user accuracy ranged from 95% to 100% with a mean of 97%, and producer accuracy ranged from 94% to 100%, with a mean of 99%. *S. patens* user accuracy ranged from 96% to 100% with a mean of 99% and producer accuracy ranged from 96% to 100%, with a mean of 98% (Supplementary Table S1). Overall, November 2019 produced the highest classification accuracy while June 2022 produced the lowest (Supplementary Table S1). This finding is also represented in the November 2019 classification prediction map (Figure 4, Panel H) which demonstrates the least amount of variation in classification while the June prediction map (Figure 4, Panel K) shows the most amount of variation.

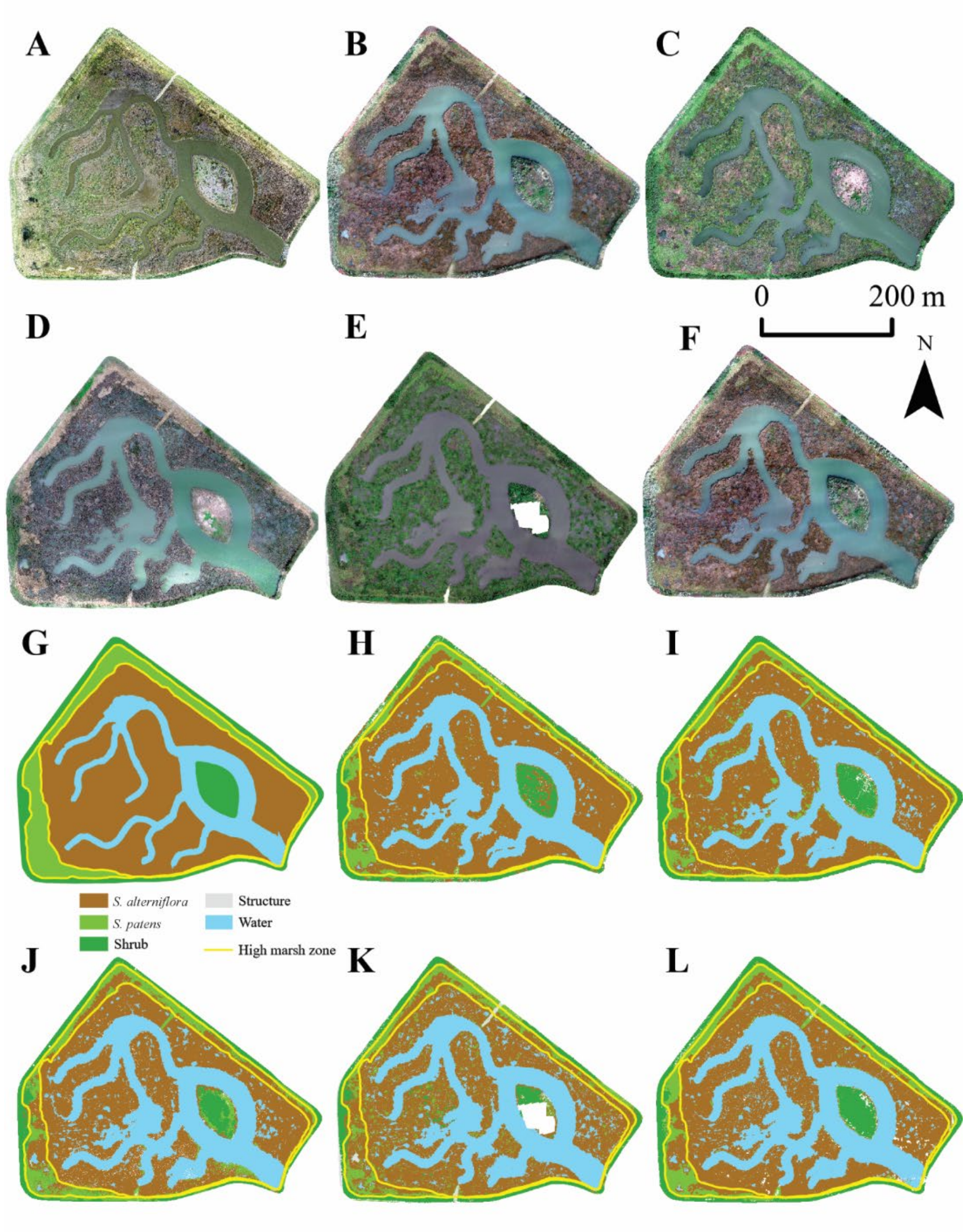


Figure 4. (A) Aerial image (1 m) of Cell 3D taken in May 2007, (B–F) UAS red-green-blue (RGB) reflectance orthomosaics (0.05 m) of Cell 3D collected in (B) November 2019, (C) August 2020, (D) April 2021, (E) June 2022, and (F) November 2022, (G) Classified map according to the 2005 as-build

planting (H–L) Random forest classification prediction maps for each UAS orthomosaic shown in panels (B–F). Yellow lines on classified maps represent the 2005 as- built planting boundaries of high and low marsh. White patch on Panel E is the result of the SfM software failing to stitch together images in that location.

Variable importance

Variable importance plots demonstrated that across all seasons, canopy elevation was the most important variable in the accuracies of the random forest classifications, except for April where texture was the most important (Figure 5). Texture was the second most important variable in November and August while red reflectance was the second most important in June. Following canopy elevation and texture, NDVI was the most important spectral variable in November and April, likely due to high contrast in the red and NIR bands from senescing vegetation. In August, the blue band was the most important spectral variable and in June, the red and red edge reflectance were the most important spectral variables which is the narrow range (600–750 nm) where (Artigas and Yang, 2005; Artigas and Yang, 2006) found separability in hyperspectral spectra of the two species. The red edge reflectance was the least important variable for November 2019, August, and April.

When excluding both texture and elevation from the classification model, overall accuracy declined an average 4%. When excluding just elevation, average overall accuracy declined 1%, and when excluding just texture, average overall accuracy stayed the same (Figure 6). Average overall accuracy of classifications applied to resampled UAS bands + NDVI were in general lower than the original classification accuracy. Average overall accuracy for 1.2, 2.4, and 4 m resolution were all 3% lower and the 10 m resolution was 7% lower average overall accuracy (Figure 6). Across all classification combinations, *S. alterniflora* user's accuracy was the lowest, followed by S. *alterniflora* producer's accuracy, demonstrating that *S. alterniflora* was more challenging to classify (Figure 6).

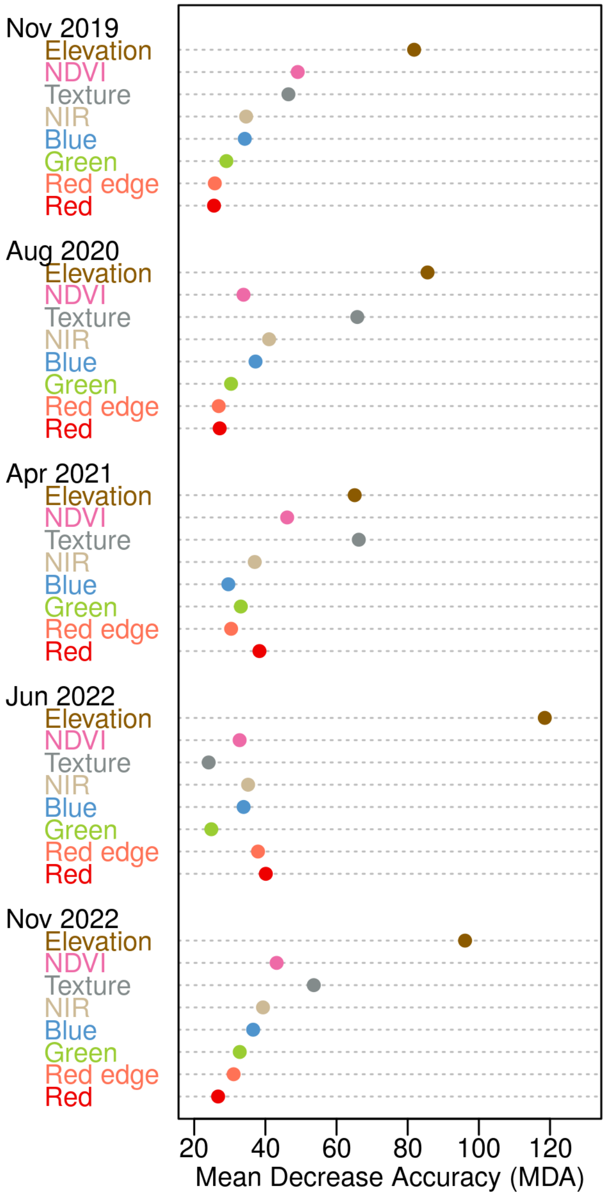


Figure 5. Variable importance plots calculated using the mean decrease accuracy (MDA) from random forest classification models generated independently for each dataset. MDA is a measure of the performance of the model without each variable, where a higher value indicates greater importance of that variable in predicting vegetation.

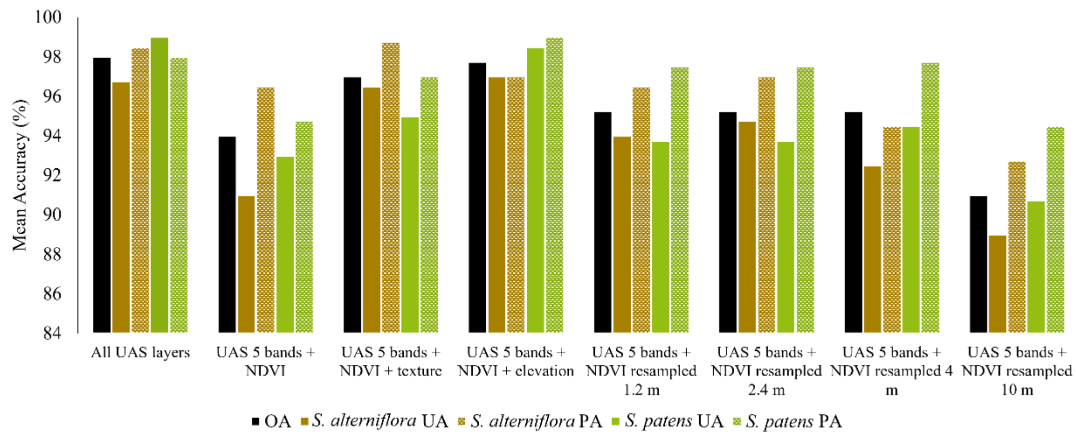


Figure 6. Bar plot of random forest classification accuracies (%) averaged across datasets when applied to various combinations of raster layers. OA = Overall accuracy, UA = User's accuracy, PA = Producer's accuracy.

Discussion

This study demonstrates how ultra-high resolution UAS remote sensing can result in high classification accuracy allowing for fine scale interannual wetland vegetation change monitoring. Supervised classification models incorporated the synergy between UAS-derived multispectral bands, canopy elevation, and texture to produce predicted habitat classes that were validated with field data. Overall, producer's and user's accuracy were evaluated between different combinations of variables and with down sampled (coarser) UAS imagery. Results have broad implications for improving remotely sensed vegetation classifications of wetlands, but several aspects and caveats of this study merit additional discussion.

Spectral and structural differentiation of S. alterniflora and S. patens

Belonging to the same genus, *S. alterniflora* and *S. patens* are easily distinguishable by the human eye, with *S. alterniflora* dominating in the low marsh growing as a tall and straight grass and *S. patens* dominating in the high marsh having a decumbent growth habit, earning it the nickname "marsh hay" (Figure 7). However, these species can be spectrally similar and oftentimes indistinguishable from a 2D aerial view. Like most vascular vegetation, *Spartina* spp. change color due to differences in light absorption and reflection properties deriving from changes in pigment concentrations throughout the growing season. In the visible spectral region, high absorption of radiation in the blue and red wavelengths is primarily due to the photosynthetic pigment, chlorophyll *a*. Absorption is not quite as strong at green wavelengths, which is why vegetation often appears green and contains a small reflectance peak around 560 nm (Figure 2). After the growing season (May–October), as both vegetation species begin to senesce, other pigments such as the carotenoids, xanthophylls, and anthocyanins also influence absorption properties (Knipling, 1970). During this time, the *Spartina* species are more easily distinguished, with *S. alterniflora* adopting a browner hue and *S. patens* remaining greener, which was observable in the November and April datasets (Figure 4).

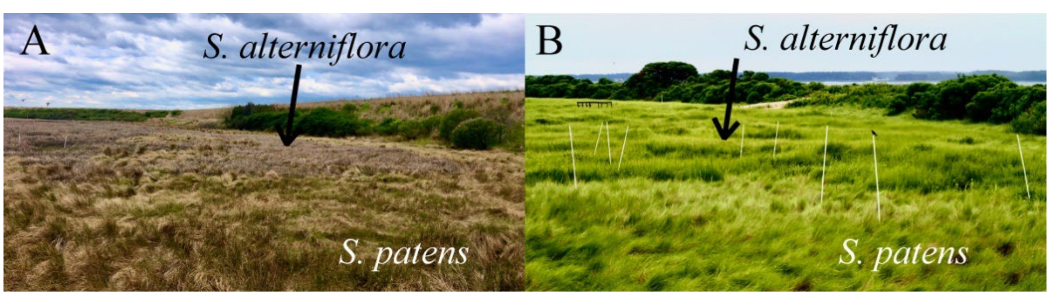


Figure 7. Field photos capturing spectral and structural differences between Spartina alterniflora and Spartina patens during the (A) fall and (B) summer seasons.

In this study, reflectance was overall lower in the blue, green, and red wavelengths (475, 560, 668 nm) compared to the red edge and NIR wavelengths (717 & 842 nm, Figure 2), similar to findings of other studies (Hardisky et al., 1983; Artigas and Yang, 2006; Kearney et al., 2009). High reflectivity in the NIR across all seasons is due to the internal cellular structure where radiation is scattered from interior parts of a leaf such as the microcavities and cellulose cell walls (Knipling, 1970). Peak NIR reflectance varies across different plant species; therefore, the ratio or slope of red and NIR wavelengths is often used for distinguishing different plant species (Knipling, 1970; Artigas and Yang, 2005; Artigas and Yang, 2006; Kearney et al., 2009; Mutanga et al., 2012). The ratio of red and NIR (NDVI) of *S. patens* was consistently higher across all seasons, except for April and June when the species appear spectrally similar during the growing season (Figure 3). The red edge slope of *S. patens* was also consistently higher than *S. alterniflora*, in contrast to other studies where a slightly higher red edge slope was associated with *S. alterniflora* (Artigas and Yang, 2006). The disparity between these two studies could be due to a number of factors, including the difference between *in situ* and UAS- derived reflectance, lighting conditions, or the time of year measurements were collected. Both *Spartina* species produced a similar spectral shape with a green reflectance peak apparent in the August and June datasets, due to stronger absorption in the blue and red wavelengths from higher chlorophyll *a* concentrations (Figure 2). However, *S. patens* reflectance was overall higher in magnitude than *S. alterniflora,* predominantly due to canopy geometry with the stems and leaves of *S. patens* laying horizontally, increasing surface area compared to upright canopy forms like *S. alterniflora,* with less surface area to capture light (Bartlett, 1981).

Studies have demonstrated how the inclusion of structural characteristics such as elevation can improve wetland vegetation classification due to distinct differences in the vertical structure of vegetation species (Maxa and Bolstad, 2009; Halls and Costin, 2016; Samiappan et al., 2017; Pricope et al., 2022). Studies have traditionally derived elevation measurements using airborne LiDAR which performs well in coastal environments like dunes (Mancini et al., 2013; Seymour et al., 2018), but tends to overestimate marsh "bare earth" elevations and vegetation height due to an inability to penetrate dense vegetation (Hladik and Alber, 2012). Many studies instead incorporate canopy elevation, which is elevation of the highest point in the point cloud, referenced to a vertical datum. Elevation from UAS-SfM produces results similar to terrestrial LiDAR and RTK GPS measurements (Mancini et al., 2013; Kalacska et al., 2017; Seymour et al., 2018), but still tends to overestimate 'bare earth' elevation from a lack of bare earth points in the SfM point cloud (Supplementary Figure S2, DiGiacomo et al., 2020). In this study, the UAS-SfM DSMs were applied in the classification models representing canopy elevation which improved classification accuracy. Canopy elevation of *S. patens* was higher than *S.*

alterniflora across all seasons, except June which could potentially be due to comparable vegetation heights between the two species during the growing season.

While reflectance spectra can describe average spectral variations of vegetation across bands, textural features can provide information about the spatial distribution of spectral variations within a band (Haralick et al., 1973). Incorporating textural features into wetland classification is a common approach since spectrally similar vegetation can exhibit distinguishable textural properties. Laba et al. (2010) and Lane et al. (2014) demonstrated higher overall accuracy of coastal wetland classification mapping when incorporating NIR-derived textural features in high resolution satellite imagery. Feng et al. (2015) also showed improved overall accuracy results when incorporating a texture analysis of 7 cm UAS imagery. The size of the moving window upon which texture features are calculated is an important consideration since a larger window can average distinct vegetation species. It is recommended to choose a window size at the optimal scale that represents the highest between-class variation and lowest within-class variation, which should be in accordance with the average size of vegetation at a study site (Feng et al., 2015).

Random forest model classification results

The classified outputs from each dataset contained high overall accuracy with values ranging from 97% to 99% (Figure 6; Supplementary Table S1). The November 2022 dataset had the lowest overall accuracy (97%) stemming from low producer's accuracy for each class (80%–100%). Low producer's accuracy is equivalent to high omission errors, which indicates that confirmed vegetation classes were not correctly predicted in the classified output. In this dataset, the “structure” class (e.g., roads, docks) had the lowest producer’s accuracy (80%), which could potentially be due to high glare present in the imagery being misclassified as high reflecting structure. The November 2022 imagery was collected the earliest of all flights (10: 00 local time); therefore, it is suggested to fly near solar noon to eliminate misclassifications of glare. In contrast, the November 2019 dataset had the highest overall accuracy (99%) and producer’s accuracy of *S. alterniflora* and *S. patens* (100%), likely due to the significant contrasting spectral differences that appear after *S. alterniflora* senescence. Producer’s accuracy of the “structure” class was still low (80%), which could be misclassification of glare or bright reflectance from bare Earth (such as in the habitat island, Figure 4).

One of the main objectives of this study included identifying which variables were most important to make accurate predictions of wetland vegetation. Variable importance differed between the random forest models due to differences in the spectral and structural characteristics of the marsh vegetation at the time of data collection. The overall high MDA for the canopy elevation and texture variables indicates the importance of these structural characteristics and the reduction in accuracy of classification models lacking these variables (Figure 5). Across all datasets except April, canopy elevation was the most important variable with much higher MDA, similar to findings from Pricope et al. (2022). In April, texture was slightly higher in importance presumably due to structural differences at the end of winter; the structure of *S. patens* is typically persistent while *S. alterniflora* tends to shed leaves and decomposes. Similarly, texture was also high in importance for the November and August datasets, likely due to considerable structural differences during and right after the growing season. Across seasons, MDA was lowest for the visible bands, demonstrating that using only visible bands can lead to lower classification accuracy. NDVI and the NIR band were higher in importance than the visible bands, due to distinct differences in NIR reflection between the two species (Knipling, 1970).

The high spatial resolution of the UAS imagery (0.05 m) produced higher classification accuracies than the down-sampled UAS imagery, demonstrating that higher spatial detail is important for accurate marsh vegetation classification (Figure 6). Overall accuracy decreased 3% when down-sampling to 1.2, 2.4, and 4 m and 7% when down-sampling to 10 m. This was likely due to larger pixel sizes averaging two or more classes, leading to misclassifications of the true habitat class. Supplementary Figure S3 illustrates the differences in spatial resolution between UAS and high-resolution WorldView 2 (2.4 m) and 3 (1.2 m) imagery, demonstrating how vegetation classification would be challenging and imprecise with coarser pixel sizes. Future work should consider calculating a "Pareto Boundary," a method that quantitatively analyzes the influence of low-resolution bias. This concept can provide insight into whether lower classification accuracy of a low spatial resolution map is given by poor performance of the model or by the low resolution of the remotely sensed data (Boschetti et al., 2004). The classification accuracies from this study are comparable to results from other UAS wetland vegetation classification studies (Feng et al., 2015; Samiappan et al., 2017; Abeysinghe et al., 2019) These studies, along with the present study, demonstrate that ultra-high resolution data produced by a low-altitude UAS can improve classification results and is a superior method to other remote sensing platforms. To distinguish wetland vegetation with high accuracy, it is recommended to collect data during the fall after *S. alterniflora* has senesced to obtain more pronounced spectral and structural differences between species at that time.

Marsh vegetation change at Poplar Island

Widespread marsh lateral migration has been observed along many Mid-Atlantic wetlands, with rates ranging from 0.1 mm yr^{-1} to 6.78 m yr^{-1} (Flester & Blum, 2020; Molino et al., 2021). However, these rates are not necessarily continuous and can be episodic with high interannual variability, as shown in this study. From 2005 to 2022, the high marsh zone that was planted with 100% coverage of *S. patens*, shifted to 45% *S. patens* coverage, indicating a significant vegetation change and evidence of marsh migration (Figure 4). From 2019 to 2022, interannual variability of high marsh *S. patens* coverage estimated from the classification model ranged from 42% to 56%. Interannual variation is common in tidal wetlands whose vegetation tends to stabilize through feedbacks that vary with depth and duration of tidal inundation (Kirwan and Megonigal, 2013). Morris et al. (2002) found that the growth of *S. alterniflora* is positively correlated with interannual variations in sea level. Relative annual mean sea level at Poplar Island has gradually increased since 1971 at an estimated rate of 3.9 mm yr^{-1} (NOAA Station 8571892, NOAA Tides and Currents) (Kent, 2015). While our UAS sample size is small (n = 4), the percentage of classified *S. alterniflora* in the high marsh boundary appears to correlate with annual mean sea level ($r^2 = 0.92$), with a lag of about 1 year. This finding suggests that *S. alterniflora* growth is influenced by water levels the previous year; however, more data is needed to confirm this potential lag in the vegetation response to water levels.

Marsh vegetation change can also be attributed to other factors including nutrient enrichment. Studies have shown that increased nutrient availability can change ratios of above and belowground biomass (Darby and Turner, 2008), competitive interactions (Bertness et al., 2002) and vertical accretion rates (Morris et al., 2002), potentially altering vegetation zonation, community composition, and resilience to sea level rise (Emery et al., 2001; Morris et al., 2013). The dredged material used for tidal marsh creation at Poplar Island is fine-grained and contains high levels of nitrogen (Staver et al., 2020); therefore, it is possible that detected vegetation changes are an indirect result of nutrient enrichment. Furthermore, the small interannual differences in classified marsh vegetation could be related to the classification accuracy of the random forest model developed for each year. While overall accuracy is high across all datasets, there is still potential for misclassification of vegetation, leading to small discrepancies in predicted vegetation coverage among years.

Benefits and caveats of UAS remote sensing in marsh environments

UAS remote sensing is a low cost, rapid, and repeatable method for collecting high resolution imagery that can be used to resolve fine scale changes in coastal wetlands without significant disturbance of the sensitive ecosystem (Joyce et al., 2019; Davis et al., 2022; DiGiacomo et al., 2022). UAS provides flexibility in obtaining imagery and can do so under optimal conditions which is useful for smaller scale projects that may be limited in field monitoring resources or projects where aircraft flights and access to satellite imagery are beyond the scope or budget. Consumer grade UAS can cover a relatively large area (13 ha in ~40 min), which will only improve with advancements in battery life. For assessing larger wetlands, fixed-wing UAS can cover larger areas due to increased flight efficiency or UAS data can be used for training and validating satellite-based classification models (Gray et al., 2018).

As with any developing technology, there are some caveats with UAS remote sensing that should be discussed. Operating a UAS is highly dependent upon the legal restrictions governing an area and can be prohibited in some areas. In the United States, UAS pilots are required to be licensed (i.e., obtain a Federal Aviation Administration (FAA) Part 107 remote pilot license) and follow local, state, and federal regulations on UAS operations. Furthermore, the cost-benefit of UAS technology should also be considered. Currently, UAS pricing can range from $500 (USD) for consumer-grade aircraft to $10K-$20K (USD) for professional-grade aircraft and sensor packages, with more advanced sensors (e.g., hyperspectral, LiDAR) reaching $50K–$100K (USD). Additionally, the software typically used for post processing (e.g., Pix4D, Agisoft Metashape) can add additional costs on the order of $1K–$5K. Computing power and storage is also a consideration since a large number of images to be processed requires large amounts of computer random access memory (RAM), large graphics processing units (GPUs) and extensive storage (Morgan et al., 2022).

UAS remote sensing in marsh environments also has its specific challenges. Studies have shown that vegetated areas with standing water and changing tide levels can interfere with the radiance being collected by the UAS and alter vegetation indices such as NDVI (Kearney et al., 2009; Doughty and Cavanaugh, 2019). Canopy structure can also influence classification results; with changes in vegetation orientation or geometry (e.g., vegetation folding over from wind or water) altering canopy optical properties and leading to errors in classifications (Knipling, 1970).

There are also some caveats and considerations with the classification approach used in this study. Other machine learning classification algorithms exist such as support vector machines, K-Nearest neighbors, or artificial neural networks (Dronovo et al., 2012; Berhane et al., 2018; Morgan et al., 2022) that could perform better or similar to the random forest models applied in this study. Some studies have evaluated various classification approaches and found random forest outperforming others (Berhane et al., 2018) or concluded that no single approach is consistently better and should be matched to the research or management question (Dronova et al., 2012). Some studies have also shown promising results using an object-based image analysis (OBIA) of wetland vegetation classification where images are segmented into spectrally homogeneous “objects” or groups of pixels that can be classified (Moffett and Gorelick, 2013; Dronova, 2015; Durgan et al., 2020). OBIA approaches cannot be fully automated and require user inputs to set segmentation and scale settings, which depend on the type of vegetation being classified and imagery resolution. Due to a range of wetland vegetation sizes, it can be challenging to select optimal segmentation settings for an entire wetland landscape, leading to over-segmentation and under-segmentation errors (Liu and Xia, 2010; Kim et al., 2011; Moffett and Gorelick, 2013). In contrast, several studies have found pixel-based classifications performing better than object-based approaches, demonstrating the model being able to identify smaller patches of vegetation (Abeysinghe et al., 2019).

Conclusion

This study demonstrates that high resolution UAS imagery is an effective tool for detecting fine scale changes in the distribution of wetland vegetation. The combination of multispectral reflectance, texture, and canopy elevation derived from UAS and SfM processing in a supervised random forest classification model led to high overall accuracy as compared to other studies. Results from the classification maps demonstrated a shift in species distribution affecting 45% of the high marsh zone over a ~17 years time frame; although recent distributions (past 4 years) were more stable, suggesting temporary equilibrium. The increased awareness of marsh habitat loss has prompted an era of creation and restoration of coastal wetlands to mitigate the effects of marsh habitat loss. Similar to large scale experiments, restoration projects require clear goals, monitoring, and adaptive management practices in order to understand site specific relationships between coastal wetlands and the surrounding environment (Zedler, 2000; Li et al., 2018). For decades, satellite and aircraft remote sensing have been used to assess and monitor coastal wetlands; however, spatial resolution, satellite revisit times, and the high cost of flying an aircraft can significantly limit the ability of managers to accurately detect fine scale vegetation changes. This study demonstrated that classification performance metrics are similar when UAS imagery is down-sampled to 1–2 m; but performance decreases with lower spatial resolution beyond this point. Repeatable, on-demand, ultra-high resolution UAS surveys can detect instances of vegetation change which can ultimately be used as early warning signals of marsh habitat loss and help identify areas for restoration efforts.

Supplementary material

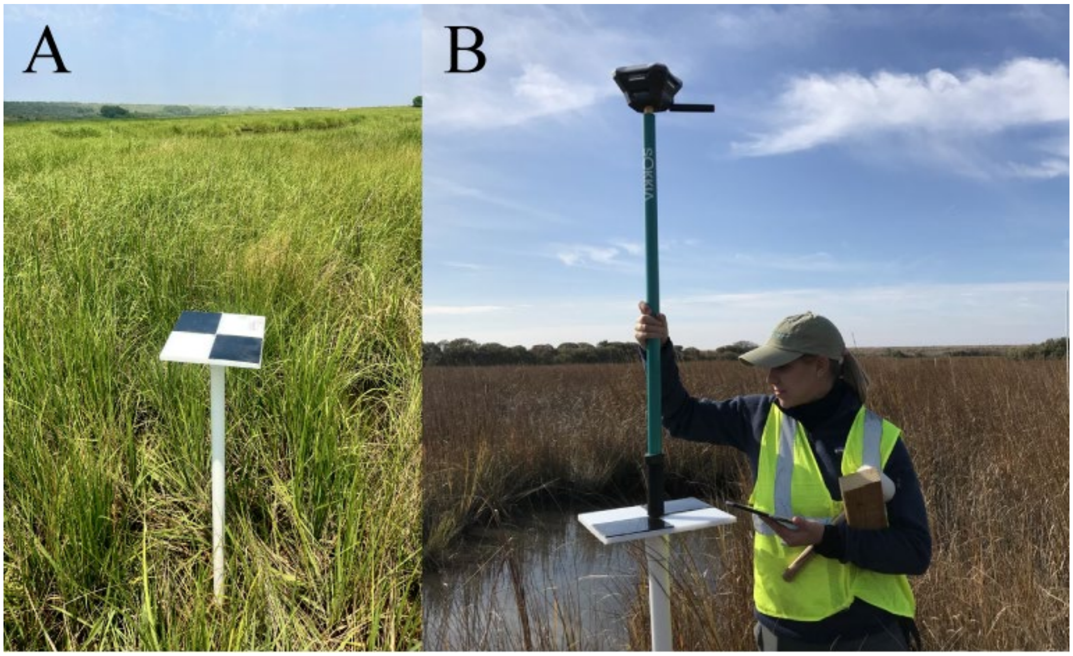


Figure S1. (A) Ground control point (GCP) deployed in the marsh and (B) surveying the GCP with an RTK GPS.

UAS-SfM elevation analysis

To explore the accuracy of UAS-SfM derived elevation in a marsh environment, measurements from the DTMs and DSMs were compared to elevation checkpoints collected with the RTK GPS. Similar elevation values between the DTM and DSM demonstrates the difficulty in separating ground from vegetation canopy points because the SfM processing is not able to produce accurate 'bare earth' terrain points due to dense vegetation (Figure S2).

Across all datasets, DTM elevation differences ranged from -0.28 to 0.78 m with a mean difference of 0.28 m while DSM elevation differences ranged from -0.11 to 0.83 m with a mean difference of 0.29 m. When compared to the RTK GPS elevation measurements, the UAS-derived DTM/DSM elevations were generally higher, apart from some elevations collected in June and August (Figure S2). This can be attributed to the DTM/DSM providing elevation values from points that are higher than bare earth since the SfM processing cannot penetrate the dense vegetation like

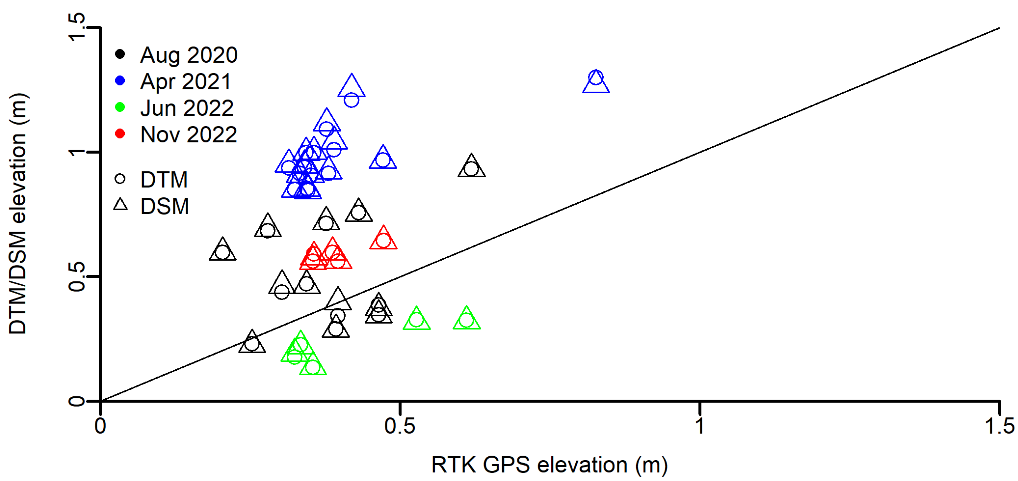


Figure S2. Scatterplot of elevation measurements extracted from the UAS-SfM derived DTM (open circles) and DSM (open triangles) at the location of RTK GPS elevation measurements. Elevation measurements are colored black, blue, green, and red for the August 2020, April 2021, June 2022, and November 2022 datasets, respectively. RTK GPS checkpoints were not collected in November 2019. Black line is the 1:1 line.

Table S1. Accuracy assessment results from a random forest classification of each dataset.

Observed data (field data)

Predicted data (map data)

November 2019

	S. alt	*S. patens*	shrub	structure	water	Row total	User's accuracy (%)
S. alt	**37**	0	0	2	0	39	95
S. patens	0	**46**	0	0	0	46	100
shrub	0	0	**36**	0	0	36	100
structure	0	0	0	**8**	0	8	100
water	0	0	0	0	**25**	25	100
Column total	37	46	36	10	25		
Producer's accuracy (%)	100	100	100	80	100		
Total correct	152						
Overall accuracy (%)	**99**						

August 2020

	S. alt	*S. patens*	shrub	structure	water	Row total	User's accuracy (%)
S. alt	**37**	2	0	0	0	39	95
S. patens	0	**44**	0	0	0	44	100
shrub	0	0	**36**	0	0	36	100
structure	0	0	0	**11**	0	11	100
water	0	0	0	0	**25**	25	100
Column total	37	46	36	11	25		
Producer's accuracy (%)	100	96	100	100	100		
Total correct	153						
Overall accuracy (%)	**99**						

April 2021

	S. alt	*S. patens*	shrub	structure	water	Row total	User's accuracy (%)
S. alt	**35**	0	0	1	0	36	97
S. patens	0	**44**	0	0	0	44	100
shrub	0	2	**33**	0	0	35	94
structure	0	0	0	**10**	0	10	100
water	0	0	0	0	**28**	28	100
Column total	35	46	33	11	28		
Producer's accuracy (%)	100	96	100	91	100		
Total correct	150						
Overall accuracy (%)	**98**						

June 2022

	S. alt	*S. patens*	shrub	structure	water	Row total	User's accuracy (%)
S. alt	**33**	0	0	0	0	33	100
S. patens	1	**44**	0	1	0	46	96
shrub	1	0	**30**	0	0	31	97
structure	0	0	0	**23**	0	23	100
water	0	0	0	0	**34**	34	100
Column total	35	44	30	24	34		
Producer's accuracy (%)	94	100	100	96	100		
Total correct	164						
Overall accuracy (%)	**98**						

November 2022

	S. alt	*S. patens*	shrub	structure	water	Row total	User's accuracy (%)
S. alt	**37**	0	0	1	0	38	97
S. patens	0	**45**	1	0	0	46	98
shrub	0	0	**37**	1	0	38	97
structure	0	1	0	**8**	0	9	89
water	0	0	0	0	**26**	26	100
Column total	37	46	38	10	26		
Producer's accuracy (%)	100	98	97	80	100		
Total correct	153						
Overall accuracy (%)	**97**						

References

Abeysinghe, T., Milas, A. S., Arend, K., Hohman, B., Reil, P., Gregory, A., et al. (2019). Mapping invasive Phragmites australis in the Old Woman Creek estuary using UAV remote sensing and machine learning classifiers. *Remote Sens.* 11 (11), 1380. doi:10. 3390/rs11111380

Aerts, R. (1999). Interspecific competition in natural plant communities: Mechanisms, trade-offs and plant-soil feedbacks. *J. Exp. Bot.* 50 (330), 29–37. doi:10.1093/jxb/50.330.29

Alizad, K., Hagen, S. C., Medeiros, S. C., Bilskie, M. v., Morris, J. T., Balthis, L., et al. (2018). Dynamic responses and implications to coastal wetlands and the

surrounding regions under sea level rise. *PloS one* 13 (10), e0205176. doi:10.1371/journal.pone. 0205176
Artigas, F. J., and Yang, J. S. (2005). Hyperspectral remote sensing of marsh species and plant vigour gradient in the New Jersey Meadowlands. *Int. J. Remote Sens.* 26 (23), 5209–5220. doi:10.1080/01431160500218952
Artigas, F. J., and Yang, J. (2006). Spectral discrimination of marsh vegetation types in the New Jersey Meadowlands, USA. *Wetlands* 26 (1), 271–277. doi:10.1672/0277- 5212
Assmann, J. J., Kerby, J. T., Cunliffe, A. M., and Myers-Smith, I. H. (2018). Vegetation monitoring using multispectral sensors—best practices and lessons learned from high latitudes. *J. Unmanned Veh. Syst.* 7 (1), 54–75. doi:10.1139/juvs- 2018-0018
Barbier, E. B., Hacker, S. D., Kennedy, C., Koch, E. W., Stier, A. C., and Silliman, B. R. (2011). The value of estuarine and coastal ecosystem services. *Ecol. Monogr.* 81 (2), 169–193. doi:10.1890/10-1510.1
Bartlett, D. S. (1981). *In situ* spectral reflectance studies of tidal wetland grasses. *Photogrammetric Eng. Remote Sens.* 47 (12), 1695–1703.
Beckett, L. H., Baldwin, A. H., and Kearney, M. S. (2016). Tidal marshes across a chesapeake bay subestuary are not keeping up with sea-level rise. *PloS one* 11 (7), e0159753. doi:10.1371/journal.pone.0159753
Belgiu, M., and Dragut, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS J. photogrammetry remote Sens.* 114, 24–31. doi:10.1016/j.isprsjprs.2016.01.011
Berhane, T. M., Lane, C. R., Wu, Q., Autrey, B. C., Anenkhonov, O. A., Chepinoga, V. V., et al. (2018). Decision-tree, rule-based, and random forest classification of high- resolution multispectral imagery for wetland mapping and inventory. *Remote Sens.* 10 (4), 580. doi:10.3390/rs10040580
Bertness, M. D., Ewanchuk, P. J., and Silliman, B. R. (2002). Anthropogenic modification of New England salt marsh landscapes. *Proc. Natl. Acad. Sci.* 99 (3), 1395–1398. doi:10.1073/pnas.022447299
Bertness, M. D. (1991). Zonation of spartina patens and spartina alterniflora in new england salt marsh. *Ecology* 72 (1), 138–148. doi:10.2307/1938909
Boschetti, L., Flasse, S. P., and Brivio, P. A. (2004). Analysis of the conflict between omission and commission in low spatial resolution dichotomic thematic products: The Pareto Boundary. *Remote Sens. Environ.* 91 (3-4), 280–292. doi:10.1016/j.rse.2004.02.015
Breiman, L. (2001). Random forests. 45(1), 5, doi:10.1023/a:1010933404324
Byrd, K. B., O'Connell, J. L., di Tommaso, S., and Kelly, M. (2014). Evaluation of sensor types and environmental controls on mapping biomass of coastal marsh emergent vegetation. *Remote Sens. Environ.* 149, 166–180. doi:10.1016/j.rse.2014.04.003

Darby, F. A., and Turner, R. E. (2008). Below- and aboveground biomass of Spartina alterniflora: Response to nutrient addition in a Louisiana salt marsh. *Estuaries Coasts* 31, 326–334. doi:10.1007/s12237-008-9037-8

Davis, J., Giannelli, R., Falvo, C., Puckett, B., Ridge, J., and Smith, E. (2022). Best practices for incorporating UAS image collection into wetland monitoring efforts: A guide for entry level users. *NOAA Tech. Memo. Nos. NCCOS* 308, 26. pps. doi:10.25923/ ccvg-ze70

DiGiacomo, A. E., Bird, C. N., Pan, V. G., Dobroski, K., Atkins-Davis, C., Johnston, D. W., et al. (2020). Modeling salt marsh vegetation height using unoccupied aircraft systems and structure from motion. *Remote Sens.* 12 (14), 2333. doi:10.3390/rs12142333

DiGiacomo, A. E., Giannelli, R., Puckett, B., Smith, E., Ridge, J. T., and Davis, J. (2022). Considerations and tradeoffs of UAS-based coastal wetland monitoring in the Southeastern United States. *Front. Remote Sens.* 3, 81. doi:10.3389/frsen.2022.924969

Doughty, C. L., and Cavanaugh, K. C. (2019). Mapping coastal wetland biomass from high resolution unmanned aerial vehicle (UAV) imagery. *Remote Sens.* 11 (5), 540. doi:10.3390/rs11050540

Dronova, I., Gong, P., Clinton, N. E., Wang, L., Fu, W., Qi, S., et al. (2012). Landscape analysis of wetland plant functional types: The effects of image segmentation scale, vegetation classes and classification methods. *Remote Sens. Environ.* 127, 357–369. doi:10.1016/j.rse.2012.09.018

Dronova, I. (2015). Object-based image analysis in wetland research: A review. *Remote Sens.* 7 (5), 6380–6413. doi:10.3390/rs70506380

Duarte, C. M., Losada, I. J., Hendriks, I. E., Mazarrasa, I., and Marbà, N. (2013). The role of coastal plant communities for climate change mitigation and adaptation. *Nat. Clim. change* 3 (11), 961–968. doi:10.1038/nclimate1970

Durgan, S. D., Zhang, C., Duecaster, A., Fourney, F., and Su, H. (2020). Unmanned aircraft system photogrammetry for mapping diverse vegetation species in a heterogeneous coastal wetland. *Wetlands* 40 (6), 2621–2633. doi:10.1007/S13157-020-01373-7

Elmore, A. J., Engelhardt, K. A. M., Cadol, D., and Palinkas, C. M. (2016). Spatial patterns of plant litter in a tidal freshwater marsh and implications for marsh persistence. *Ecol. Appl.* 26 (3), 846–860. doi:10.1890/14-1970

Emery, N. C., Ewanchuk, P. J., and Bertness, M. D. (2001). Competition and salt-marsh plant zonation: Stress tolerators may be dominant competitors. *Ecology* 82 (9), 2471–2485. doi:10.1890/0012-9658(2001)082[2471:casmpz]2.0.co;2

Enwright, N. M., Griffith, K. T., and Osland, M. J. (2016). Barriers to and opportunities for landward migration of coastal wetlands with sea-level rise. *Front. Ecol. Environ.* 14 (6), 307–316. doi:10.1002/fee.1282

Evans, T. L., Costa, M., Tomas, W. M., and Camilo, A. R. (2014). Large-scale habitat mapping of the Brazilian pantanal wetland: A synthetic aperture radar approach. *Remote Sens. Environ.* 155, 89–108. doi:10.1016/j.rse.2013.08.051

Feagin, R. A., Luisa Martinez, M., Mendoza-Gonzalez, G., and Costanza, R. (2010). Salt marsh zonal migration and ecosystem service change in response to global sea level rise: A case study from an urban region. *Ecol. Soc.* 15 (4), art14. doi:10.5751/es-03724- 150414

Feng, Q., Liu, J., and Gong, J. (2015). UAV Remote sensing for urban vegetation mapping using random forest and texture analysis. *Remote Sens.* 7 (1), 1074–1094. doi:10.3390/rs70101074

Flester, J. A., and Blum, L. K. (2020). Rates of mainland marsh migration into uplands and seaward edge erosion are explained by geomorphic type of salt marsh in Virginia coastal lagoons. *Wetlands* 40 (6), 1703–1715. doi:10.1007/s13157-020-01390-6

Gray, P. C., Ridge, J. T., Poulin, S. K., Seymour, A. C., Schwantes, A. M., Swenson, J. J., et al. (2018). Integrating drone imagery into high resolution satellite remote sensing assessments of estuarine environments. *Remote Sens.* 10 (8), 1257. doi:10.3390/ rs10081257

Gross, M. F., Hardisky, M. A., Klemas, V., and Wolf, P. L. (1987). Quantification of biomass of the marsh grass spartina alterniflora Loisel using landsat thematic mapper imagery. *Photogrammetric Eng. Remote Sens.* 53 (11), 1577–1583.

Halls, J., and Costin, K. (2016). Submerged and emergent land cover and bathymetric mapping of estuarine habitats using worldView-2 and LiDAR imagery. *Remote Sens.* 8 (9), 718. doi:10.3390/rs8090718

Haralick, R. M., Shanmugam, K., and Dinstein, I. H. (1973). Textural features for image classification. *IEEE Trans. Syst. man, Cybern.* (6), 610–621. doi:10.1109/tsmc. 1973.4309314

Hardisky, M. A., Smart, R. M., and Klemas, V. (1983). Growth response and spectral characteristics of a short Spartina alterniflora salt marsh irrigated with freshwater and sewage effluent. *Remote Sens. Environ.* 13 (1), 57–67. doi:10.1016/0034-4257(83)90027-5

Hladik, C., and Alber, M. (2012). Accuracy assessment and correction of a LIDAR-derived salt marsh digital elevation model. *Remote Sens. Environ.* 121, 224–235. doi:10. 1016/j.rse.2012.01.018

Howes, N. C., Fitzgerald, D. M., Hughes, Z. J., Georgiou, Y., Kulp, M. A., Miner, M. D., et al. (2010). Hurricane-induced failure of low salinity wetlands. *Proc. Natl. Acad. Sci.* 107 (32), 14014–14019. doi:10.1073/pnas.0914582107

Joyce, K. E., Duce, S., Leahy, S. M., Leon, J., and Maier, S. W. (2019). Principles and practice of acquiring drone-based image data in marine environments. *Mar. Freshw. Res.* 70 (7), 952–963. doi:10.1071/MF17380

Kalacska, M., Chmura, G. L., Lucanus, O., Bérubé, D., and Arroyo-Mora, J. P. (2017). Structure from motion will revolutionize analyses of tidal wetland landscapes. *Remote Sens. Environ.* 199, 14–24. doi:10.1016/j.rse.2017.06.023

Kearney, M. S., Stutzer, D., Turpie, K., and Stevenson, J. C. (2009). The effects of tidal inundation on the reflectance characteristics of coastal marsh vegetation. *J. Coast. Res.* 25 (6), 1177–1186. doi:10.2112/08-1080.1

Kent, J. (2015). Water level variations at poplar island. MD: NOAA Technical Report NOS CO-OPS 076.

Kim, M., Warner, T. A., Madden, M., and Atkinson, D. S. (2011). Multi-scale GEOBIA with very high spatial resolution digital aerial imagery: Scale, texture and image objects. *Int. J. Remote Sens.* 32 (10), 2825–2850. doi:10.1080/ 01431161003745608

Kirwan, M. L., and Megonigal, J. P. (2013). Tidal wetland stability in the face of human impacts and sea-level rise. *Nature* 504 (7478), 53–60. doi:10.1038/nature12856

Kirwan, M. L., Walters, D. C., Reay, W. G., and Carr, J. A. (2016). Sea level driven marsh expansion in a coupled model of marsh erosion and migration. *Geophys. Res. Lett.* 43 (9), 4366–4373. doi:10.1002/2016GL068507

Klemas, V. (2013). Remote sensing of coastal wetland biomass: An overview. *J. Coast. Res.* 29 (5), 1016–1028. doi:10.2112/JCOASTRES-D-12-00237.1

Knipling, E. B. (1970). Physical and physiological basis for the reflectance of visible and near-infrared radiation from vegetation. *Remote Sens. Environ.* 1 (3), 155–159. doi:10.1016/s0034-4257(70)80021-9

Laba, M., Blair, B., Downs, R., Monger, B., Philpot, W., Smith, S., et al. (2010). Use of textural measurements to map invasive wetland plants in the Hudson River National Estuarine Research Reserve with IKONOS satellite imagery. *Remote Sens. Environ.* 114 (4), 876–886. doi:10.1016/j.rse.2009.12.002

Lane, C. R., Liu, H., Autrey, B. C., Anenkhonov, O. A., Chepinoga, V. V., and Wu, Q. (2014). Improved wetland classification using eight-band high resolution satellite imagery and a hybrid approach. *Remote Sens.* 6 (12), 12187–12216. doi:10.3390/ rs61212187

Li, X., Bellerby, R., Craft, C., and Widney, S. E. (2018). Coastal wetland loss, consequences, and challenges for restoration. *Anthr. Coasts* 1 (1), 1–15. doi:10.1139/ anc-2017-0001

Liu, D., and Xia, F. (2010). Assessing object-based classification: Advantages and limitations. *Remote Sens. Lett.* 1 (4), 187–194. doi:10.1080/ 01431161003743173

Liu, M., Li, H., Li, L., Man, W., Jia, M., Wang, Z., et al. (2017). Monitoring the invasion of Spartina alterniflora using multi-source high-resolution imagery in the Zhangjiang Estuary, China. *Remote Sens.* 9 (6), 539. doi:10.3390/rs9060539

Mahdavi, S., Salehi, B., Granger, J., Amani, M., Brisco, B., and Huang, W. (2018). Remote sensing for wetland classification: A comprehensive review. *GIScience Remote Sens.* 55 (5), 623–658. doi:10.1080/15481603.2017.1419602

Mancini, F., Dubbini, M., Gattelli, M., Stecchi, F., Fabbri, S., and Gabbianelli, G. (2013). Using unmanned aerial vehicles (UAV) for high-resolution reconstruction of topography: The structure from motion approach on coastal environments. *Remote Sens.* 5 (12), 6880–6898. doi:10.3390/rs5126880

Massetti, A., Sequeira, M. M., Pupo, A., Figueiredo, A., Guiomar, N., and Gil, A. (2016). Assessing the effectiveness of RapidEye multispectral imagery for vegetation mapping in Madeira Island (Portugal). *Eur. J. Remote Sens.* 49 (1), 643–672. doi:10. 5721/EuJRS20164934
Maxa, M., and Bolstad, P. (2009). Mapping northern wetlands with high resolution satellite images and LiDAR. *Wetlands* 29 (1), 248–260. doi:10.1672/08-91.1
Mendelssohn, I. A., Mckee, K. L., and Patrick, W. H. (1981). Oxygen deficiency in spartina alterniflora roots: Metabolic adaptation to anoxia. *Science* 214 (4519), 439–441. doi:10.1126/science.214.4519.439
Moffett, K. B., and Gorelick, S. M. (2013). Distinguishing wetland vegetation and channel features with object-based image segmentation. *Int. J. Remote Sens.* 34 (4), 1332–1354. doi:10.1080/01431161.2012.718463
Molino, G. D., Defne, Z., Aretxabaleta, A. L., Ganju, N. K., and Carr, J. A. (2021). Quantifying slopes as a driver of forest to marsh conversion using geospatial techniques: Application to chesapeake bay coastal-plain, United States. *Front. Environ. Sci.* 9, 149. doi:10.3389/fenvs.2021.616319
Morgan, G. R., Hodgson, M. E., Wang, C., and Schill, S. R. (2022). Unmanned aerial remote sensing of coastal vegetation: A review. *Ann. GIS* 2, 1–15. doi:10.1080/19475683. 2022.2026476
Morris, J. T., Shaffer, G. P., and Nyman, J. A. (2013). Brinson review: Perspectives on the influence of nutrients on the sustainability of coastal wetlands. *Wetlands* 33 (6), 975–988. doi:10.1007/s13157-013-0480-3
Morris, J. T., Sundareshwar, P. v., Nietch, C. T., Kjerfve, B., and Cahoon, D. R. (2002). Responses of coastal wetlands to rising sea level. *Ecology* 83 (10), 2869–2877. doi:10. 1890/0012-9658(2002)083[2869:ROCWTR]2.0.CO;2
Mutanga, O., Adam, E., and Cho, M. A. (2012). High density biomass estimation for wetland vegetation using WorldView-2 imagery and random forest regression algorithm. *Int. J. Appl. Earth Observation Geoinformation* 18, 399–406. doi:10.1016/ j.jag.2012.03.012
Nardin, W., Taddia, Y., Quitadamo, M., Vona, I., Corbau, C., Franchi, G., et al. (2021). Seasonality and characterization mapping of restored tidal marsh by NDVI imageries coupling UAVs and multispectral camera. *Remote Sens.* 13 (21), 4207. doi:10.3390/ rs13214207
Pendleton, L., Donato, D. C., Murray, B. C., Crooks, S., Jenkins, W. A., and Sifleet, S. (2012). Estimating global "blue carbon" emissions from conversion and degradation of vegetated coastal ecosystems. *PLoS ONE* 7, e43542. doi:10.1371/ journal.pone.0043542
Pinton, D., Canestrelli, A., Wilkinson, B., Ifju, P., and Ortega, A. (2020). A new algorithm for estimating ground elevation and vegetation characteristics in coastal salt marshes from high-resolution UAV-based LiDAR point clouds. *Earth Surf. Process. Landforms* 45 (14), 3687–3701. doi:10.1002/esp.4992

Pricope, N. G., Minei, A., Halls, J. N., Chen, C., and Wang, Y. (2022). UAS hyperspatial LiDAR data performance in delineation and classification across a gradient of wetland types. *Drones* 6, 268. doi:10.3390/drones6100268
Qi, M., MacGregor, J., and Gedan, K. (2020). Biogeomorphic patterns emerge with pond expansion in deteriorating marshes affected by relative sea level rise. *Limnol. Oceanogr.* 66 (4), 1036–1049. doi:10.1002/lno.11661
Ridge, J. T., and Johnston, D. W. (2020). Unoccupied aircraft systems (UAS) for marine ecosystem restoration. *Front. Mar. Sci.* 7, 438. doi:10.3389/fmars.2020. 00438
Samiappan, S., Turnage, G., Allen Hathcock, L., and Moorhead, R. (2017). Mapping of invasive phragmites (common reed) in Gulf of Mexico coastal wetlands using multispectral imagery and small unmanned aerial systems. *Int. J. Remote Sens.* 38 (8-10), 2861–2882. doi:10.1080/01431161.2016.1271480
Schieder, N. W., Walters, D. C., and Kirwan, M. L. (2018). Massive upland to wetland conversion compensated for historical marsh loss in chesapeake bay, USA. *Estuaries Coasts* 41, 940–951. doi:10.1007/s12237-017-0336-9
Seymour, A. C., Ridge, J. T., Rodriguez, A. B., Newton, E., Dale, J., and Johnston, D. W. (2018). Deploying fixed wing unoccupied aerial systems (UAS) for coastal morphology assessment and management. *J. Coast. Res.* 34 (3), 704–717. doi:10.2112/ JCOASTRES-D-17-00088.1
Silva, T. S. F., Costa, M. P., and Melack, J. M. (2010). Spatial and temporal variability of macrophyte cover and productivity in the eastern amazon floodplain: A remote sensing approach. *Remote Sens. Environ.* 114 (9), 1998–2010. doi:10.1016/j.rse.2010.04.007
Staver, L. W., Stevenson, J. C., Cornwell, J. C., Nidzieko, N. J., Staver, & K. W., Owens,
M. S., et al. (2020). Tidal marsh restoration at poplar island: Ii. Elevation trends, vegetation development, and carbon dynamics. *Wetlands* 40 (6), 1687–1701. doi:10. 1007/s13157-020-01295-4/
Story, M., and Congalton, R. G. (1986). Accuracy assessment: A user's perspective. *Photogrammetric Eng. remote Sens.* 52 (3), 397–399.
USACE (2020). *Multispectral imagery analysis for vegetative coverage monitoring at poplar island*. Baltimore, Maryland: USACE.
Wan, H., Wang, Q., Jiang, D., Fu, J., Yang, Y., and Liu, X. (2014). Monitoring the invasion of Spartina alterniflora using very high resolution unmanned aerial vehicle imagery in Beihai, Guangxi (China). *Sci. World J.* 2014, 1–7. doi:10. 1155/2014/638296
Wang, M., Fei, X., Zhang, Y., Chen, Z., Wang, X., Tsou, J. Y., et al. (2018). Assessing texture features to classify coastal wetland vegetation from high spatial resolution imagery using Completed Local Binary Patterns (CLBP). *Remote Sens.* 10 (5), 778. doi:10.3390/rs10050778
Westoby, M. J., Brasington, J., Glasser, N. F., Hambrey, M. J., and Reynolds, J. M. (2012). 'Structure-from-Motion' photogrammetry: A low-cost, effective tool

for geoscience applications. *Geomorphology* 179, 300–314. doi:10.1016/j.geomorph.2012.08.021
Zedler, J. B. (2000). Progress in wetland restoration ecology. *Trends Ecol. Evol.* 15 (10), 402–407. doi:10.1016/s0169-5347(00)01959-5

Chapter 6: Conclusions

At the interface between terrestrial environments and open oceans, coastal waters encompass many unique habitats, provide a suite of ecosystem benefits, and support a growing proportion of the world's population. Coastal ecosystems are sensitive areas responding at various scales (events to long-term) where monitoring and management can be challenging. Recent progress in sensor radiometric, spatial, spectral, and temporal resolutions, new data processing methods, and advancing technology make environmental remote sensing an adaptable and useful tool to study and monitor coastal areas. This dissertation includes four research chapters demonstrating how satellite and UAS remote sensing can be incorporated to enhance the study of, monitoring, and management of the Chesapeake Bay estuarine environment.

Satellite remote sensing

Satellite remote sensing provides water quality observations at spatial and temporal scales that exceed current field monitoring techniques and can be used to enhance the understanding of coastal processes and monitoring of marine resources. However, at the top of the atmosphere, satellite measurements include both the surface under observation and the intervening atmosphere. Since water-leaving radiance is at most 10% of the total radiance a satellite measures, accurate atmospheric correction is essential. The challenge of atmospheric correction is exacerbated in coastal waters due to the optically complex water-leaving reflectance signal and the presence of strongly absorbing atmospheric aerosols with differing compositions and vertical distributions. Chapter 2 evaluates four different atmospheric correction algorithms applied to the Ocean Land and Color Instrument (OLCI) onboard the Copernicus Sentinel-3A and -3B satellites. OLCI was chosen for this study since it can currently resolve optical data in coastal water bodies at higher spatial (300 m), spectral (21 bands), and temporal (2-day) resolutions than any other operational ocean color satellite sensor. A matchup analysis with a regional *in situ* dataset from Chesapeake Bay waters demonstrated that the Case 2 Regional Coast Color (C2RCC) atmospheric correction processor had the best performance, particularly in the longer wavelengths. The results from this study agreed and contrasted with similar studies in regional coastal water bodies, depicting the challenge of developing a robust atmospheric correction algorithm that works in all global coastal waters. Future research will include comparing biophysical products derived from C2RCC processed data to an extensive dataset of *in situ* water quality concentrations collected in the Chesapeake Bay to evaluate performance and analyze water quality variability. This study provides a framework with associated uncertainties and recommendations to utilize OLCI ocean color data to monitor biogeochemical dynamics in the Chesapeake Bay.

Remote sensing reflectance (R_{rs}), the product of atmospheric correction processors, is governed by the type and concentration of optically active constituents (OACs), namely phytoplankton pigments, chromophoric dissolved organic matter, inorganic sediment, and water itself. Therefore, R_{rs} can be used to infer underlying water quality metrics. In coastal waters where OACs are dynamic and vary independently from each other, retrieving water quality concentrations remains a challenge. In Chapter 3, a machine learning classification of normalized OLCI R_{rs} was used to identify fifteen dominant optical water types (OWTs, i.e., color) of the Chesapeake Bay. A comprehensive matchup analysis with *in situ* water quality data characterized each OWT allowing for comparisons across space and time. The accuracy of seven empirical algorithms designed to estimate concentrations of chlorophyll *a* from measurements of R_{rs} was analyzed for each OWT. The best performing algorithm was selected and future work will include applying these algorithms in a dynamic switching algorithm to produce an improved blended chlorophyll *a* retrieval of the Chesapeake Bay. This study was the first of its kind to apply an OWT classification to study spatial and temporal dynamics of ecologically and biogeochemically important constituents in the Chesapeake Bay and has broad implications for water quality monitoring and management.

UAS remote sensing

Traditional *in situ* based methods of collecting data in the field or by boat can be inadequate in space and time, expensive, time-consuming, and potentially harmful to the environment. Traditional satellite remote sensing can capture variability throughout time and space; however, limitations including the presence of clouds, atmospheric effects, land adjacency effects, and spatial resolution can hinder effective monitoring. UAS fills an operational gap between *in situ* and satellite remote sensing methods and can provide the spatial and temporal variability required for useful monitoring in a dynamic and rapidly evolving environment.

While the currently available commercial multispectral UAS sensor technology is geared toward terrestrial applications, mostly precision agriculture, the spectral bands can be used to retrieve water quality parameters in aquatic water bodies. Atmospheric effects can usually be ignored in low altitude UAS flights; however, the effect of sun glint and surface reflected light need to be accounted for in order to obtain the highest accuracy of water quality data. The inclusion of surface reflected light can lead to an overestimation of R_{rs} and remotely sensed water quality retrievals. In Chapter 4, R_{rs} was measured from the MicaSense RedEdge-MX multispectral sensor onboard a consumer-grade UAS. This study compared four approaches to remove sun glint and surface reflected light and found that a pixel-based deglinting procedure utilizing the brightness of the NIR band performed best when compared to *in situ* R_{rs} measurements. Future work will include improving water quality retrieval algorithms to resolve small-scale phenomena and physical processes such as patchy algal blooms and turbulence characteristics.

UAS remote sensing is an adaptable tool to monitor coastal wetland environments at much greater spatial resolution than sensors on aircraft or satellites. Resultant imagery can also be readily rendered in three dimensions through Structure from Motion (SfM) photogrammetric processing. Chapter 5 demonstrated how UAS remote sensing can serve as an effective tool for detecting fine scale changes in the distribution of wetland vegetation, using an engineered wetland at Poplar Island, Chesapeake Bay as a case study. A combination of multispectral reflectance, texture, and canopy elevation derived from UAS and SfM processing in a supervised random forest classification model led to high overall accuracy as compared to other studies. Results from the classification maps revealed a long-term shift in species distribution. This study demonstrates how repeatable, on-demand, ultra-high resolution UAS surveys can detect instances of vegetation change which can be used as early warning signals of marsh habitat loss or marsh migration and help identify areas for restoration efforts.

Selection of monitoring method

For the past 40+ years, satellite remote sensing has described spatial dynamics of coastal environments across multiple spatial scales. Satellites collect near-real time, global data in large swaths, providing a means to study an entire environment at once, such as the entire Chesapeake Bay estuary. However, satellites remain incapable of resolving small-scale events and can miss rapidly evolving phenomena that often occur in coastal environments. As discussed in detail in this dissertation, satellite remote sensing retrievals are also dependent on the accuracy of atmospheric correction procedures.

UAS observe fine scale resolutions and can be deployed at frequencies that fill the satellite observational gap. UAS are appealing due to their modular sensing capabilities and on-demand deployment and can carry sensors that collect data in similar spectral resolutions to a satellite. However, UAS cover only a small fraction of what a satellite covers, limiting the extent of the monitoring area. Additionally, UAS have some methodological challenges including sun glint and reflected skylight removal and maintaining appropriate viewing geometries.

In situ data will always be critical to collect in order to validate remote sensing data. It is considered the most accurate data since it involves directly accessing the environment to deploy a measuring instrument or collect a sample from. However, *in situ* data are limited by spatial coverage, fieldwork frequency, and methods changes, and can rapidly change by dynamic coastal processes, which can hinder the ability to gain a comprehensive understanding of the spatial and temporal patterns in coastal environments.

When combined, *in situ*, UAS, and satellite remote sensing methods can fill observational gaps. A synergy of these data streams can more accurately capture spatiotemporal complexity across orders of magnitude, advancing the study of, monitoring, and management of dynamic coastal environments.