Dear Sirs:

Thank you for the opportunity to participate in the *Biogeosciences* review process. The paper for which I am the corresponding author is titled, "A Global End-Member Approach to Derive $a_{\mathrm{CDOM}}(440)$ from Near-Surface Optical Measurements" (BG-2019-259) by Stanford Hooker, Atsushi Matsuoka, Raphael Kudela, Youhei Yamashita, Koji Suzuki, and Henry Houskeeper. In the material presented below, page and line numbers are abbreviated as capital single letters, e.g., P2 L1–5 refers to page 2 lines 1 through 5. The reviewers comments are presented first followed by subsequent editor comments.

Reviewer 1 Comment 1: The reviewer believes a better justification is needed for the high vertical resolution even for turbid waters and questions if the pressure sensor used in the C-OPS achieve such a resolution.

Authors Response 1: The accuracy of the pressure transducers used in this study have a depth resolution, in terms of precision, of 0.03–0.08 mm in all water masses. The manuscript will be modified by adding the precision values to the "less than 1 mm" vertical resolution statements in the abstract and P6 L25.

Reviewer 1 Comment 2: The reviewer states some content is redundant and believes the authors spend much effort describing the global perspective of sampling sites (P2 L27–28 and P3–4), which the reviewer believes is a simple concept that can be properly combined and shortened.

Authors Response 2: While the authors agree that maintaining brevity is crucial, defining the usage of a "global perspective" is necessary given the unique diversity of water masses sampled in this study, and because the two most common definitions are based on spatial extent and dynamic range. For example, if an algorithm is only accurate in the open ocean, such an algorithm would apply to the vast majority of the pixels in a worldwide remote sensing image and, thus, could be considered *global* even though it provides degraded—perhaps useless—information in the coastal zone and inland waters. The authors set out to create an algorithm that could be applied to the open ocean, coastal zone, and inland waters with equal efficacy, i.e., an algorithm that was arguably global in its application in terms of the dynamic range of water masses rather than the spatial extent of water masses.

Not wanting to belabor possible inadequacies in prior word usage, perhaps based primarily on spatial extent, the authors kept this topic short by *not* introducing *regional* and *universal* perspectives, which could be considered applicable to algorithm development. Reviewer 2, however, provided comments regarding the universal application of the algorithm, wherein a purported goal of an algorithm is to observe the environment "as is." Consequently, this material will be modified in the manuscript in keeping with the Authors Response 24 to Reviewer 2 while maintaining brevity to the greatest extent practicable.

Reviewer 1 Comment 3: The reviewer suggests combining the descriptions on the algorithms developed in Hooker et al. (2013) in P2 L1–9 and P3 L11–18 to avoid repetition and make it more clear to readers.

Authors Response 3: The authors agree and will move the applicable material presented in P3 to P2 within the manuscript.

Reviewer 1 Comment 4: The reviewer questions whether the statement that "this study includes coastal and inland ecosystems that are typically too small to be studied using common remote sensing platforms" (P2 L17–18) is accurate, because the reviewer asserts "there are many new sensors with high spatial resolution that have been used to study inland and estuarine waters" (e.g., MSI, OLCI, and OLI).

Authors Response 4: The intention of the authors was not to dismiss the fairly large body of work on coastal waters using MSI and OLI (among others), but rather to point out that many of the field sites included in this analysis are still not effectively sampled by those platforms, because three contiguous water pixels are required to ensure that imagery is free from edge effects (i.e., stray light). This limits sensors such as MSI, OLI, and OLCI to water bodies with spatial scales exceeding 30, 90, and 900 m, and this study sampled water masses that were smaller in one or more of these dimensions (e.g., rivers). While the aforementioned sensors have been used extensively for coastal and inland waters, e.g., Palmer et al. (2015) and Mouw et al. (2015), high spatial resolution sensors typically have reduced spectral resolution or range, and existing methods for characterizing $a_{\mathrm{CDOM}}(\lambda)$ have been inadequate even within relatively large, lacustrine water bodies (Kutser et al. 2015). The manuscript will be modified to briefly clarify this point (P2 L17–18).

Reviewer 1 Comment 5: The reviewer recommends combining the descriptions on the algorithms developed in Hooker et al. (2013) in P2 L1–9 and P3 L11–18 to avoid repetition and make it more clear to readers.

Authors Response 5: The authors agree and will modify the manuscript by moving the applicable material presented in P3 to P2.

Reviewer 1 Comment 6: The reviewer believes it is better to move the summarized importance of $a_{\mathrm{CDOM}}(440)$ in P2 L19–25 to immediately after the sentence ending in P2 L3.

Authors Response 6: The authors agree and will modify the manuscript accordingly.

Reviewer 1 Comment 7: The reviewer suggests shortening the descriptions on global perspective sampling sites to avoid repetition.

Authors Response 7: The authors agree and will modify the manuscript (P7–8) accordingly.

Reviewer 1 Comment 8: The reviewer suggests using a detailed number to replace the phrase "almost all" in P7 L9.

Authors Response 8: The authors agree and will modify the manuscript by replacing "Almost all" with "Approximately 98%" (P7 L9).

Reviewer 1 Comment 9: The reviewer suggests adding the time period over which the field dataset was collected (P8 L13).

Authors Response 9: The authors agree and will modify the manuscript by adding "data collection spanned 29 April 2013 to 25 January 2017" (P8 L13)).

Reviewer 1 Comment 10: The reviewer suggests adding a reference regarding the conversion to absorption coefficient (P11 L9).

Authors Response 10: The authors agree and will modify the manuscript by adding the Green and Blough (1994) citation in P11 L9.

Reviewer 1 Comment 11: The reviewer suggests more quantitative information of some of the criteria used for subcategories is needed in Sect. 2.5, such as, what's the chlorophyll *a* value used to define algal bloom and what's the dominant species of HAB? The reviewer notes that the authors may want to mention this information as some of the HABs, like red tide species and cyanobacteria blooms display totally different optical properties, and further distinct Kd spectra.

Authors Response 11: Defining a HAB event is somewhat fuzzy. For example, Smayda (1997) (correctly) points out that an absolute chlorophyll concentration has little to do with how so-called "blooms" are labeled, while HABs can occur at very low chlorophyll if they are dominated by noxious or toxic algae. The authors have followed typical convention for identification of a HAB as an event dominated by a known HAB organism, or causing a deleterious effect on humans or the environment. The authors will expand the manuscript slightly (P12 L26–27) to indicate that a HAB is subjectively determined with these objective criteria.

Reviewer 1 Comment 12: The reviewer notes that the authors mentioned that a sample sometimes satisfied more than one subcategory, and then requests additional information as to why the authors classified data in this way which uses little quantitative information since the authors already utilized K-mean classification of Kd spectra, which make more sense.

Authors Response 12: The subjective classification scheme makes use of relevant factual information known about a water mass, which is not always available in algorithm development or validation, thus the inclusion of both objective and subjective approaches in this study. The text introducing the subjective subcategories (P23 L14–21) will be modified to clarify the underlying characteristics of each subcategory are based on factual and quantitative observations essential to the qualities and attributes of each one, i.e., the subcategories are not arbitrary or lacking in substance.

As presented at the start of Sect. 2.5, the subjective subcategories were based on sampling information directly associated with optical properties, which would otherwise not be known. The subcategories also provided a running inventory of the types of water masses that were incrementally sampled to ensure a global sampling with a representative number of each type of water mass to the extent practicable. This approach was effective as proved by the amount of data in each FCM class (Table 2): N_1 244, N_2 263, N_3 305, N_4 265, and N_5 94. The amounts are rather balanced except for the most extreme observations in the N_5 class, which were the most difficult to obtain (extreme water masses are not common).

The fewer samples in the extreme N_5 class do not negate the overall applicability of the subcategorization scheme. For each FCM class, the cumulative percentage of the dynamic range in K_d end members and $a_{\mathrm{CDOM}}(440)$ is, respectively, as follows: N_1 2.3% and 3.4%; N_2 8.4% and 8.9%; N_3 22.7% and 36.4%; N_4 82.9% and 100.0%; and N_5 100.0% and 100.0%. These data reveal that the $a_{\mathrm{CDOM}}(440)$ dynamic range is completely established with the addition of the N_4 class, which provides the largest extension of the optical and biogeochemical dynamic ranges. The N_1 (case-1) class makes the smallest contribution to the dynamic ranges, although it arguably accounts for most of the pixels in a global CDOM image, and the N_5 class only extends the optical dynamic range. Consequently, the global algorithm is comprehensively established by the N_1–N_4 combined classes and the N_5 class primarily contributes variance to the algorithm, i.e., it does not expand the optical versus biogeochemical relationship provided by classes N_1–N_4.

An algorithm based on all the data in classes N_1–N_4 yields a linear fit of $y = 0.2317x - 0.0053$, the RMSE is 5.3%, and the slope is to within 9.3% of the original value presented by Hooker et al. (2013). This result is significantly similar (the slopes agree to within 2.7%) to the subjective results discussed as the "fourth more comprehensive data set" in Sect. 4 (P29 L19–23), wherein $y = 0.2379x - 0.0049$, the RMSE is 6.2%, and the slope is to within 6.9% of the original value presented by Hooker et al. (2013). Consequently, the robustness of the algorithm is directly supported by the combination of subjective and objective classifications, with the latter using fuzzy c-means (FCM). It is important to remember that the FCM approach is different from a hard or crisp classification, such as k-means which was not used.

The authors believe the subjective-objective approach that was used facilitates an appropriate interpretation of the results, because the subjective approach reveals cause-and-effect relationships (e.g., the importance of water masses subjected to resuspension effects in higher class numbers as shown in Table 2), and the objective approach provides an unbiased strictly quantitative confirmation. Table 2 will be expanded a little bit to present the extent of the dynamic range analysis and summarize its importance. In addition, the manuscript will be modified to succinctly present the N_1–N_4 algorithm results in companion with the "fourth more comprehensive data set" as a fifth comprehensive data set in Sect. 4 (P29 after L23).

Reviewer 1 Comment 13: The reviewer suggests a comparison of the Kd ratio with aCDOM(440) for each K-mean classified cluster to see if the correlations can be improved for each group.

Authors Response 13: The authors agree that comparisons within each FCM category would be valuable, and the extra material will be succinctly presented in the manuscript Sect. 4 and linked with the new material to be added as part of Authors Response 41 (which will appear at the end of Sect. 3.7 P27 L13, i.e., right before Sect. 4 (P27 L14).

Reviewer 1 Comment 14: The reviewer strongly recommends the authors add one more section in the Results section to display some of interesting dataset (e.g., Kd and/or Ed spectra) collected in conservative and non-conservative water masses, such as, Hypersaline Lakes, river mouth, HAB.

Authors Response 14: While the authors appreciate the interest in K_d and E_d spectra, they must respectfully decline for the following reasons: a) the objective of the paper is to produce a simple reliable algorithm using end-member analyses and not to describe the optical properties of a large number of water masses; b) the study contains so many observations that individual spectra overlap each other significantly and create a continuum of lines from the pure water limit to the most turbid N_5 water body (White Lake), which are difficult to discern and which do not provide any additional information beyond what is already presented in the manuscript; c) if multiple panels are used to magnify certain parts of the dynamic range, the amount of required text significantly lengthens the manuscript; and d) there is persistent pressure from both reviewers to produce a more compact manuscript. Consequently, no new figures will be added to the manuscript and the only new sections are the result of moving material at the request of the reviewers (see Authors Responses 19 and 20).

Reviewer 1 Comment 15: The reviewer notes the authors displayed the correlations between Kd ratio and aCDOM(440) for different categories, however, they did not evaluate the performance of the algorithms, it's better to keep some dataset for validating their algorithm and study the errors and uncertainties using statistic parameters like RMSE, ARE, R2.

Authors Response 15: The approach of the original Hooker et al. (2013) manuscript, as well as this study, already adhere to the Reviewer's comment as follows: a) Hooker et al. (2013) used a separate validation subset collected in significantly different geographical areas, which was used to confirm the efficacy of the derived algorithm and quantify performance parameters (e.g., RMSE, R^2, etc.); b) the entire dataset used for this study is a distinct and separate set of observations from significantly different geographical areas that is used to validate the original algorithm; c) the categorization scheme partitions the observations into conservative and non-conservative water masses and the influence of adding in the non-conservative fraction is quantified in terms of the same performance parameters, so there is a progression of algorithm validations based on increasingly larger data subsets; and d) NOMAD data were the only independent validation subset that could be found, but the poorer spectral and geographical diversity limits their utility. To clarify this aspect of validation, the manuscript will be slightly modified in the new NOMAD Sect. 2.6 (see Authors Response 19) to remind the reader that the validation process used herein adheres to the concept of having a separate dataset to estimate the uncertainties using statistical parameters and the use of the NOMAD data is an extension of that philosophy.

Reviewer 1 Comment 16: The reviewer notes the authors mentioned better correlation for hypersaline or alkaline lakes compared to the overfilled lakes, and explained turbidity could be the possible disturbance (P16 L3–7), and suggested the following: a) more information should be provided, such as, how the turbid water modified the spectrum of Rrs, Ed and further Kd; b) what type of sediment, like mud, clay or silty increased the turbidity; and c) identification of the "atypical constituents" in L7, perhaps with a supporting reference; d) an explanation of how this constituent influences the Kd spectra; and e) for section 3.2, an explanation about the influence of sediment-resuspension on high turbidity and on the variations of Kd.

Authors Response 16: The referenced pages deal with lacustrine subcategories, wherein the manuscript text notes that refilled lakes frequently exhibit larger anomalies with respect to the algorithm than hypersaline or alkaline lakes, especially in terms of turbidity as determined by the K_d ratio. The "atypical constituents" introduced to a water mass when it is overfilled is a generalized phenomenon, wherein land that is subjected to other purposes (e.g., agricultural and anthropogenic activities associated with grazing, farming, vehicular traffic, etc.) will provide one or more constituents to the water mass when the lake overfills that are not typical of what is in the water mass prior to overfilling, because these activities are not possible in the water mass. The manuscript will be slightly modified to make this point clearer in P16 L6–7.

In regards to the list of requested clarifications the principal two problems are as follows: a) the manuscript does not rely on $R_{\mathrm{rs}}(\lambda)$ or $E_d(\lambda)$, so these variables are not a part of this study; and b) determining how turbidity modified the $K_d(\lambda)$ spectrum, determining what type of sediment increased the turbidity, identifying "atypical constituents" introduced into a lake when it overfills, explaining how the introduction of an "atypical constituent" influences K_d spectra, and explaining more about the influence of sediment resuspension on high turbidity and on the variations of $K_d(\lambda)$ all require baseline data for the subject water masses prior to modification, which are not available. Although this entire line of inquiry indicates interest in the work that the authors performed, it is not objectively focused on algorithm validation. Instead, it involves scientific pursuits that either cannot be answered, because of a lack of baseline data, or are outside the scope of the material presented (e.g., R_{rs} and E_d spectra). Consequently, the manuscript will not be modified for these itemized comments.

Reviewer 1 Comment 17: The reviewer requests more information regarding the atypical algal bloom (P17 L19–23).

Authors Response 17: The language regarding an "atypical bloom" was meant to describe a generic case of a water mass wherein there was unusually high biomass of (typically) a single species of algae. An example was given of physical forcing (wind and waves) accumulating unusually high biomass on one side of a lake, although similar phenomena occur in the coastal ocean, whereby a combination of algal growth, vertical migration (behavior), and physical aggregation can on occasion result in dinoflagellate blooms reaching more than 1,000 $\mu g\,L^{-1}$ of chlorophyll in Monterey Bay, which is about 10 times higher than (already high-biomass) red tide events that are not physically aggregated. To put this into context, the manuscript will be modified slightly by adding the Kudela et al. (2015) reference to P17 L23. The Kudela et al. (2015) study documented concentrations of chlorophyll in excess of 2,000 $\mu g\,L^{-1}$ at Pinto Lake, one of the water bodies included in Fig. 6 as anomalous.

Reviewer 1 Comment 18: The reviewer notes that the authors mentioned UV attenuation (P18 L21–22), which is likely due to production of UV-absorbing pigments (e.g., Mycosporine-like Amino Acids (MAAs)) by phytoplankton in response to UV stress, and suggested more information and a reference. The reviewer also suggested the authors may want to add information of the dominant species of algal bloom, because there are some species that can also strongly modify the spectrum in 700-800 nm range, like Trichodesmium.

Authors Response 18: The reviewer is correct, some of the blooms (marine dinoflagellates in particular) are associated with MAA-like compounds that strongly impact UV absorption. The HAB events identified in Fig. 6 as Monterey Bay were dominated by *Cochlodinium* and *Akashiwo*. For *Akashiwo* in particular, MAA-like compounds are a diagnostic indicator of the presence of foam-producing substances (Jessup et al. 2009), while Kwon et al. (2018) demonstrate significant increases in FDOM and DOC in *Cochlodinium* blooms.

Unfortunately, even if phytoplankton absorption coefficient spectra were available, it would still be difficult to unequivocally ascertain the presence of MAAs, because no valid beta factor that can rigorously be applied for the UV spectral domain is presently available. To avoid speculation, the authors decided not to expand the discussion about MAAs in the text, but the manuscript will be modified slightly with a small revision to include the aforementioned references in P18 L22.

Reviewer 1 Comment 19: The reviewer suggests the description on NASA NOMAD data should be moved to Data and Method section.

Authors Response 19: The authors agree, so the manuscript will be modified by adding a new Sect. 2.6 using the appropriate material in Sect. 3.6.

Reviewer 1 Comment 20: In regards to Sect. 3.7, the reviewer suggests some of contents relevant to method of K-mean classification seems to fit better in Methods (Sect. 2) and proposes it's better to move the whole section 3.7 up to the first sub-section in Results, which could help the general readers to better understand the algorithm performance (or nonperformance) of non-conservative waters.

Authors Response 20: Recalling that this study does not use k-means classification (it uses FCM classification), the authors agree to modify the manuscript by moving some of the FCM contents in Sect. 3.7 to a new Sect. 2.7. The authors disagree with moving Sect. 3.7 to the first section in Results, because this would place the material being moved before the description of the subjective classification data and the material being moved references these data, so the subjective classification data must appear first, so there will be no modification of the manuscript for this part of the comment.

Reviewer 2 Comment 1: The reviewer asserts there is a need to a) clearly link the subjective classification of environments to other literature discussing optical variability, and b) pre-screen for unique optical water types to effectively retrieve inherent optical properties from a given system.

Authors Response 21: The algorithm development approach espoused in the manuscript classified the data subjectively, so more complex water masses could be added to the algorithm incrementally to provide quantitative assessments of how the more complex water masses influenced algorithm performance. No other purpose was espoused or documented in the manuscript. To clarify the purpose of the subjective approach, the text introducing the subjective categories (P23 L14–21) will be modified to make it clear that the "subcategories are used exclusively to assess algorithm performance as more complex water masses are included."

The second assertion requiring a need to pre-screen for unique optical water types to effectively retrieve IOPs from a given system is not scientifically objective, because the study makes no effort to do this; in fact, it seeks to accomplish the opposite, i.e., retrieve an in-water biogeochemical constituent, $a_{\mathrm{CDOM}}(440)$, from observations of the diffuse attenuation coefficient, $K_d(\lambda)$. No modifications of the manuscript are made for this comment. Consequently, no modifications will be made to the manuscript for this comment.

Reviewer 2 Comment 2: The reviewer states the authors would be well-served to present the full data set to display dynamic range across unique environments and graphically present the ability to measure radiometric variability at the millimeter scale—a fascinating accomplishment. A clearer link between this capability and the decision to treat certain environments as anomalous is also warranted.

Authors Response 22: Figures 3–8 and 10 present all of the data used in the study while identifying all of the unique environments that were sampled with almost all of them labeled. From the perspective of the capabilities of a COTS pressure transducer, the ability to measure radiometric variability at the millimeter scale is well established. What makes the accomplishment unique is the use of small digital thrusters to maneuver the optical backplane while maintaining the planar orientation of the radiometer apertures. The latter ensures very little data, and typically no data in inland waters where vertical resolution is critical, are lost because the vertical tilt of the instruments exceed 5°. Consequently, the primary reason radiometric variability is measured at the millimeter scale is because of the the precision of the pressure transducer, the high data rate of C-OPS microradiometers, the slow sinking rate near the surface from the *hydrobaric* buoyancy chamber, plus the planar stability from the C-PrOPS accessory. The former is now included in the manuscript as part of the Authors Response to Reviewer 1 Comment 1, the latter is documented in the manuscript as part of describing the efficiencies of thruster-assisted profiling (P7 L16–22), and the data rate with Morrow et al. (2010) reference will be added as a small addition to the manuscript at P7 L16–22. The hydrobaric buoyancy chamber is documented in Fig. 1. A graphical depiction is not deemed necessary, because it will lengthen the manuscript without adding any value beyond what is already reported in the manuscript, so the manuscript will not be further modified.

Reviewer 2 Comment 3: The reviewer posits that it seems the author's treated any data that did not conform to the algorithm as anomalous or atypical.

Authors Response 23: The definition of conservative water masses (P3 L28–29) and the follow-on definition of what is considered an anomalous condition (P3 L29 to P4 L1–3) were used to establish the subjective subcategories. The merits of classifying the data using subjective criteria is easily discerned by comparing Figs. 3 and 5, i.e., conservative water masses versus a subset of non-conservative water bodies. Rivers were arguably one of the most difficult ecosystems sampled, because they are by definition rather shallow and the moving water makes profiling challenging. Nonetheless, Fig. 3 contains many examples of the inland portion of rivers, which are characterized as conservative water masses despite the sampling difficulties. The Sacramento River at flood stage, however, was categorized as a resuspension water body because resuspended material was visible, and its similarity with other resuspended water masses in respect to the algorithm is apparent in Fig. 5. River mouths (which represent a mixing of water masses) also group together, although differently than rivers or resuspension in respect to the algorithm. The merits of classifying the data was also proven by using an objective FCM scheme to show the data naturally classify into groups, and five classes were identified.

The fact that the various groups of data have dissimilar relationships with the algorithm does not mean they were treated differently. Ultimately, the manuscript shows in Sect. 4 P29 L19–23 that if all the data except extreme lacustrine water bodies (e.g., the White Lake data had estimated values in the UV domain, Bear Lake is a unique scattering anomaly created by calcium carbonate particles, etc.) are used to create a fourth data set with 1,086 observations—i.e., almost 90% of the 1,230 maximum and 93% of the data used in Table 2 to create the five objective FCM classifications—the linear fit of the fourth more comprehensive data set is $y = 0.2379x - 0.0049$, the RMSE is 6.2%, and the new slope is to within 6.9% of the original value presented by Hooker et al. (2013).

To prevent a similar erroneous understanding, the manuscript will be modified immediately after the sentence ending on P4 L25 by adding the following: "Ultimately, all subcategories are incrementally added to the algorithm evaluation process to assess performance as a function of increasing water mass complexity." Furthermore, to clarify that the subjective approach correctly categorizes the data, the manuscript will be modified slightly in P16 after L19 by adding a new sentence: "The dissimilar expression of the flooded Sacramento River with respect to the inland riverine data in Fig. 3 not in flood conditions (i.e., as conservative water masses), shows the subjective classification approach has merit."

Reviewer 2 Comment 4: The reviewer states the goal of algorithms is to observe the environment "as is", so such a subjective treatment of any water that does not conform to anticipated algorithm output doesn't seem to be appropriate.

Authors Response 24: All of the water masses sampled in this study were observed "as is" and ultimately 90% or more of the data were used to evaluate algorithm performance. The small amount of data that were not included in the fourth algorithm data set were properly excluded as detailed in the manuscript (e.g., the White Lake data had estimated values in the UV domain, Bear Lake is a unique scattering anomaly created by calcium carbonate particles, etc.).

More importantly, the analysis presented was not aimed at removing "anomalous" water masses that did not conform to the proposed end-member analysis (EMA), but rather to identify what situations lead to departures from the algorithm. The authors agree that an algorithm based on first principals should be able to fit all naturally (and artificial) occurring samples, but the authors also never claimed that there is a fundamental law or physical concept that would fit such an algorithm. Rather, the authors point out that "anomalous" water bodies that deviate from the predicted relationship can generally be explained and subjectively or objectively classified based on their optical properties.

The anomalous points may not be ideal for validation of the algorithm but are nonetheless incrementally included in evaluating algorithm performance, as presented in Sect. 4. Within the calibration validation research (CVR) paradigm, these same points would certainly be in-scope for research, and the EMA approach would help to define the research (i.e., determining what characteristics about the water masses makes them anomalous when using the EMA approach). It is important to recall that in terms of surface area, the "anomalous" water masses constitute a tiny fraction of the total aquatic area of worldwide ecosystems, thereby confirming that for the vast majority of cases the approach will work well. The manuscript will be modified at the introduction of the subjective subcategories in Sect. 2.5 (P13 L18) to clarify the "anomalous" water masses are a tiny fraction of the total aquatic area of worldwide ecosystems.

Reviewer 2 Comment 5: The reviewer asserts discussions of the "parent water mass" suggest that harbors, creek/river inputs, etc. are oddities; in fact, these spatial gradients are what we are trying to retrieve accurately with algorithms, and was a highlight of why the C-OPS was such an important instrument in the coastal zone in Hooker et al. (2013).

Authors Response 25: While the authors agree that the C-OPS instrument suite has improved characterization of coastal zone features such as spatial gradients, the parent water mass discussion is solely applied to demonstrate sensitivity, and all parent water mass modifier data were included in algorithm performance evaluations. In addition, no water masses described in the manuscript are considered to be oddities by the authors. The use of parent water mass modifiers was explained in the manuscript as "a localized alteration of water properties, e.g., a creek inflow into a lake, and demonstrates the sensitivity of the methods used herein to distinguish small changes." All the parent water mass modifier data were used to demonstrate sensitivity and all were included in algorithm performance evaluations. Furthermore, the spatial gradients associated with 90% or more of the water masses were included in evaluating algorithm performance. The small amount of data that were not included in the fourth algorithm data set were properly excluded as detailed in the manuscript (e.g., the White Lake data had estimated values in the UV domain, Bear Lake is a unique scattering anomaly created by calcium carbonate particles, etc.). The authors believe the present form of the manuscript properly presents all forms of sensitivity arguments, so no modifications will be made.

Reviewer 2 Comment 6: The reviewer thinks the author's would be best served by exploring the data, presenting the dynamic range (with associated categories, if necessary) and relate to algorithm performance.

Authors Response 26: The authors agree that presenting dynamic range and relating associated categories to algorithm performance are important, and believe that Figs. 3–8 and 10 with accompanying text succinctly and thoroughly addresses these requests. In addition, Sect. 4 already explores algorithm performance as a function of increasing water mass complexity, so no modifications to the manuscript were deemed necessary. Consequently, no modifications to the manuscript will be made.

Reviewer 2 Comment 7: The reviewer suggests discussion of potential improvements is also warranted, particularly considering that upcoming sensors are expected to have advanced spectral capabilities that, hopefully, will make band ratios less relevant for estimating a final product.

Authors Response 27: The authors have expanded on this point at the end of Sect. 4. It is important to note that while some remote sensors (e.g., PACE) will provide much improved spectral capabilities and the ability to employ other algorithm approaches (spectral shape, semi-analytical inversion, etc.) other sensors (e.g., MSI, OLI, OLCI, etc.) are still multispectral, while SBG has yet to be fully defined. Even with spectrometer-based systems (PACE, possibly SBG) the EMA approach provides a simple and independent way to assess data quality and algorithm performance for more sophisticated algorithms, while the ability to estimate CDOM at the millimeter depth scale with C-OPS, or more generally to conduct the same measurements with a two-channel system, provides an ability to generate vast quantities of data compared to traditional optical measurements, improving both validation and research for coupled remote sensing and *in situ* studies.

The manuscript will be modified after P31 L29 by adding the following: "While planned high spectral resolution sensors, such as the Plankton, Aerosol, Cloud, ocean Ecosystem (PACE) and Surface Biology and Geology (SBG) missions, may support more sophisticated retrievals of parameters such as CDOM, the simplified approach provided by end-member analysis can in principal be used with both legacy and next-generation sensors, thereby providing continuity in space and time, as well as a capability to generate high-quality in-water data with a simplified measurement approach (assuming rigorous adherence to the sampling protocols)."

Reviewer 2 Comment 8: The reviewer suggests presenting noise in measurements, and the ability to observe millimeter scale variability, would be quite useful as an additional figure.

Authors Response 28: Appropriate noise estimates already appear in the manuscript either in terms of method performance in Sect. 4 (P30 L27 to P31 L5) and in Figs. 3–8, which include all optical casts associated with each biogeochemical measurement, so the noise is represented graphically. The sensitivity of the methods used is presented graphically as part of showing the parent water mass modifiers in Fig. 7. A separate graphical depiction of millimeter scale variability is unnecessary, because it will lengthen the manuscript without adding any value beyond the metrics already reported in the manuscript. Consequently, no modifications to the manuscript were deemed necessary.

Reviewer 2 Comment 9: The reviewer suggests that the existing figures showing performance across subjectively classified environments could be reduced into subplots of a single figure, allowing for fuller presentation of the dataset, including CDOM spectra and the ability to observe such fine scale variability in radiometry.

Authors Response 29: Figures 3–8 and 10 present all of the data used in the study while identifying all of the unique environments that were sampled with almost all of them labeled. CDOM spectra were not used in the study, so their presentation is unnecessary. The ability to observe fine-scale variability is addressed in Authors Response 22.

Reviewer 2 Comment 10: The reviewer questions whether there was any consideration of using CDOM spectral slope as a proxy of conservative versus non-conservative water masses.

Authors Response 30: Although an interesting idea, the authors did not consider any alternative to identifying conservative versus non-conservative water masses for a few reasons. First, CDOM was measured separately in the various laboratories, and spectral slope was also determined using slightly different methodologies, so an analysis as to whether that had an impact would be needed. Second, our primary focus was on using the bio-optical data, with subjective or objective (clustering) criteria, and there is no simple way to derive the spectral slope from the EMA approach (it would have to be tested and validated independently of the existing methodology). Third, some authors have suggested that spectral slope does not vary as widely as is assumed, and that much of the difference is related to analytical methodology (which brings us back to the first point), e.g., Twardowski et al. 2004. Consequently, no modifications to the manuscript were considered necessary.

Reviewer 2 Comment 11: The reviewer states that a discussion of "anomalous" features is certainly warranted for radiometry measurements; however, it seems throughout that the author's measured "anomalous" conditions with the assumption that the Kd(320)/Kd(780) relationship observed for conservative waters holds to a universal truth across environments. This seems to be a gross simplification of the complexity of radiometry and a weakness of the current manuscript. As stated by the authors: P27 L16-19: "The validation approach was based on the concept that water masses evolving conservatively (i.e., free from stressors that might cause anomalies to the natural range in the gradient of a constituent) are suitable for validating the original Hooker et al. (2013) inversion algorithm for deriving aCDOM(440) from Kd($\breve{\lambda}$) spectral end members." Effectively, observations that did not conform to the algorithm are considered "anomalous" and subjectively classified.

Authors Response 31: Nowhere in the manuscript do the authors state or imply that the $K_d(320)/K_d(780)$ relationship observed for conservative waters holds to a universal truth across environments. Instead, the authors established a plausible hypothesis for establishing a starting point for the validation process, evaluated algorithm performance while incrementally adding water masses of increasing optical complexity, and ultimately established the original algorithm had an accuracy of approximately 6% while using 90% or more of the water masses spanning three decades of optical and biogeochemical dynamic range. The small amount of data that were not included in the fourth algorithm data set were properly excluded as detailed in the manuscript (e.g., the White Lake data had estimated values in the UV domain, Bear Lake is a unique scattering anomaly created by calcium carbonate particles, etc.).

The authors note that the reviewer's interpretation (that observations that did not conform to the algorithm are considered "anomalous" and subjectively classified) is not in agreement with the methodology performed in this study. In fact, the subjective classification took place prior to sampling a particular site for a particular water mass to ensure that the process was unbiased and free of any influence from the actual observations. The manuscript will be slightly modified by clarifying that the subjective classification took place prior to sampling a particular site for a particular water mass in P12 L5–7.

Reviewer 2 Comment 12: The reviewer questions why the authors did not use an objective clustering approach to classify different environments, and consider an effective algorithm for each of these water types.

Authors Response 32: The goal of the study was to produce a global algorithm that would be evaluated by adding in incrementally more complex waters while quantifying the effect on the algorithm. At no point was a series of regional or classification algorithms considered. Within this context the "anomalous" water masses are atypical compared to what would occur with conservative mixing of defined end members and have an unknown additional optical complexity. As noted in Authors Response 21, the authors do not state that the algorithm is based on some *universal truth*, and did not mean to imply that water masses that do not fit the conservative concept are not natural in the broader sense.

The argument for subjectively classifying them is because a priori subjective information was used first that is not strictly derived from the optics. For example, if optical sampling is conducted near an ice edge, or in a drained and refilled lake, then it is logical to assume that this forcing has some impact on the optical properties, but a purely objective classification scheme may not partition those conditions, because there is not necessarily a unique optical property that is associated with those physical, biological, or chemical phenomena. So the point was to start with a subjective classification scheme and identify data that do not conform to the algorithm (anomalies), then to determine whether there is some clustering based on additional information that exposes where those observations fall when compared to the expected relationship, and then to evaluate the effect on algorithm performance as these optically more complex data are incrementally added to the algorithm.

The reviewer states that observations that did not conform are anomalous, and therefore subjectively classified. All data were first subjectively classified (prior to data collection in the field, per Authors Response

31) and then plotted against the algorithm. Some did not conform to the generalized relationship and they were evaluated with additional (non-optical) information to identify a proximate cause. An objective clustering would still highlight those as non-conforming; the subjective clustering attempts to provide a rational explanation for why they differ. Separate algorithms based on the objective classification were not developed, because the goal was to test a globally-applicable algorithm and not to develop a set of related but different regional algorithms. While the latter certainly could be done, it was not the primary focus of this study.

To ensure this question is addressed in the manuscript, the manuscript will be slightly modified in P12 L5–7 by clarifying that the algorithm validation process begins with the conservative water mass data and algorithm performance is further quantified by incrementally adding more complex water masses from the subcategories to the evaluation data set.

Reviewer 2 Comment 13: The reviewer states the introduction is fascinating and an interesting discussion. However, consider that the last ¸1 page of text does not have a single citation. It seems much more relevant to tie this work more clearly into existing literature considering optical water variability due to different environments and optical complexity within a specific environment. This would leave much of the introduction intact while more clearly linking to how this builds upon past efforts, which it certainly does.

Authors Response 33: The authors agree and have added the following citations to provide the requested links: a) P3 L23 (Yapiyev et al. 2017); b) P3 L25 (Bodaker et al. 2010); c) P4 L10 (Lee and Hu 2006); d) P4 L15 (Morel 1974); e) P4 L24 (Guarch-Ribot and Butturini 2016 and Vazquez et al. 2011).

Reviewer 2 Comment 14: The reviewer states P7 Lines 16-22 are helpful and that all preceding paragraphs of this section (2.1) seem more suitable for supplementary material. The reviewer acknowledges the motivation to show and describe improvements to the instrument from that shown in Hooker et al. (2013), but the reviewer believes this takes valuable space away from a more complete presentation of the dataset. Currently, all that is shown is the instrument, sampling locations and relationships with the empirical algorithm. Displaying the dynamic range of the data would be particularly useful.

Authors Response 34: While the authors support the goal of increasing the brevity of the text, they believe that the information described is necessary given that knowledge of the technology to observe optical variability at the 1 mm scale is a) not widely known throughout the community of practice, and b) was central in enabling the accomplishments of this manuscript. In addition, the authors note that the manuscript already contains a complete presentation of the data used as well as the dynamic range of the data in Figs. 3–8 and 10. Consequently, no modifications will be made to the manuscript.

Reviewer 2 Comment 15: The reviewer questions why only the Pacific Ocean and half of the Arctic Ocean samples were baseline-corrected and wonders if use of 590-600 nm result in a significant offset for these spectra (P11 L8).

Authors Response 35: The other half of Arctic Ocean samples were also baseline-corrected with the mean value of $a_{\mathrm{CDOM}}(\lambda)$ between 683 and 687 nm according to a reference cited in the manuscript (Matsuoka et al. 2017). The effect of the three different laboratory methods (including the different wavelengths used for baseline correction) on $a_{\mathrm{CDOM}}(440)$ values were tested. As a result, it was found that the use of three different laboratory methods to determine $a_{\mathrm{CDOM}}(440)$ does not significantly influence the results presented in the manuscript. This indicates that use of 590–600 nm did not result in a significant offset for these spectra, as represented by $a_{\mathrm{CDOM}}(440)$, compared to the other methods. This issue was presented in Sect. 4 (P30 L20-29 to P31 L1-5). The manuscript will be briefly modified in P11 L8 to make it clear that the other half of the Arctic Ocean samples were baseline-corrected and also after P31 L5 to state the use of 590–600 nm did not result in a significant offset.

Reviewer 2 Comment 16: The reviewer says the section on western US coastal and inland water CDOM analysis is not clear (P11 L18–27), as follows: a) the two references to quantifying CDOM as the absorption coefficient at 440 nm and use of the Single Exponential Model make it seem that only CDOM at 440 nm was measured (there was no reference to quantifying CDOM at 440 nm for the other water samples); b) the authors, however, also mention absorption spectra were measured, perhaps using Ultrapure water to dilute the signal; and c) Were the samples optically thick and needed dilution? Please clarify this section, and to the extent possible, condense the sections on analysis of differently sourced water samples.

Authors Response 36: To clarify the description about CDOM analysis for western US coastal and inland water, the manuscript will be modified as follows: "For the western US coastal and inland waters, water samples were passed through a 0.2 μm syringe filter (Whatman GD/X) and absorbance of CDOM was measured on either a Cary Varian 50 spectrophotometer using a 10 cm quartz cell or an UltraPath liquid waveguide spectrometer with 2 m path length." Samples for the UltraPath measurements were not diluted. While absorbance can be saturated in the short part of the spectrum when using 2 m path length, this issue was not observed at 440 nm. In the visible part of the spectrum, the results are better in terms of precision than using a classical method with a 10 cm cuvette.

Reviewer 2 Comment 17: The reviewer notes that for Sect. 2.5 the categorization of optical variability across water bodies is interesting, and the details are certainly relevant. The reviewer also notes that the categorizations are rather subjective, and effectively used to explain outliers in the algorithm relationship. In the present state of the manuscript, the reviewer believes this seems quite subjective. The reviewer also believes for a universal algorithm, it seems highly relevant to look for underlying means for deviations in the relationship that would aid in how the algorithm is applied. The reviewer goes on to state that effectively, the authors have categorized the environment that results in deviating optical properties that do not perform well within the algorithm; however, the authors haven't utilized the dataset to hypothesize on specific, only general, mechanisms (e.g., minerogenic content of particles and refractive index, spectrally different CDOM, dominance of phytoplankton absorption and scattering on Kd relationships rather than a relatively generic "sediment resuspension"). The author admits that while quite difficult, the level of detail for the other sections seems to warrant this consideration. The reviewer concludes that the the authors have attempted to bypass this variability by using the end-member approach, targeting the wavelengths most and least influenced by aCDOM; other optical parameters significantly impacting the signal suggests a more detailed explanation, outside of categories, is warranted.

Authors Response 37: The authors have addressed the subjective classification topic in Authors Response 12 among others, and have established revisions in order to clarify this misunderstanding.

As presented in Sect. 4, the most extensive application of the optically complex data to the algorithm results in the use of 90% or more of all the data with an accuracy of approximately 6%. It is not scientifically objective to characterize this capability as resulting in "deviating optical properties that do not perform well within the algorithm." In fact, as noted in Sect. 4, standard algorithms, some of which do not span three decades of dynamic range in optical and biogeochemical parameters, do not perform as well (P31 L9–12), and these published algorithms do not provide the specificity requested by the reviewer to explain their large inaccuracies (some of which are significantly larger).

The authors do not believe that an algorithm that spans three decades of dynamic range in both optical and biogeochemical parameters and that has an accuracy of approximately 6% when 90% or more of the optically complex data are used needs to investigate the specific mechanisms for such a small degradation in accuracy; especially when the application of that same algorithm to conservative water masses has an accuracy of approximately 1% and the uncertainty in the optical measurements is on the order of 5%. In other words, the algorithm is clearly robust and significantly more capable than any present alternative.

The robustness is further established by creating a so-called universal algorithm, which is undefined by the reviewer, but is assumed to mean that any water mass wherein an optical profiler can be deployed is expected to be part of the evaluation of the end-member approach. In this case, the universal algorithm is constructed from *all the data from all subcategories*. The linear fit of this universal data set is $y = 0.2206x + 0.0088$, the RMSE is 7.5%, and the new slope is to within 13.7% of the original value presented by Hooker et al. (2013). In other words, the universal algorithm includes water masses that would not normally be included in an algorithm—i.e.. hypersaline, alkaline, and polluted lakes—and it equals or exceeds the performance of common so-called global algorithms.

If the hypersaline, alkaline, and polluted lakes are removed from the universal algorithm, the linear fit of this sixth comprehensive data set (the fifth comprehensive data set is presented in Authors Response 12) is $y = 0.2250x + 0.0024$, the RMSE is 6.8%, and the new slope is to within 12.0% of the original value presented by Hooker et al. (2013). The robustness of the universal and sixth comprehensive data sets can be evaluated by comparing the results to the fifth comprehensive algorithm, which used data from all the N_1–N_4 classes and completely fulfilled the dynamic range in $a_{\mathrm{CDOM}}(440)$, as presented in Authors Response 12. The linear fit of the fifth comprehensive data set is $y = 0.2317x - 0.0053$, the RMSE is 5.3%, and the new slope is to within 9.3% of the original value presented by Hooker et al. (2013).

Although of general scientific interest, the level of accuracy achieved with the universal, fourth, fifth, or sixth comprehensive data sets does not warrant investigations into the influence of minerogenic content of particles and refractive index, spectrally different CDOM, dominance of phytoplankton absorption and scattering on K_d relationships or any other source of variance, because the accuracies of the universal and comprehensive algorithms do not warrant such investigations which are otherwise beyond the scope of the work described here.

Sect. 4 of the manuscript will be modified to include the algorithm performance results for the fifth, sixth, and universal data sets; the new material will be presented after the fourth data set results in P30 L1.

Reviewer 2 Comment 18: The reviewer cites P16 L6-7 regarding new acreage from overfilled likes is a source of atypical constituents, either in composition or concentration, and states, "Really, this is the challenge of creating flexible, accurate algorithms that work across a variety of water types, either due to spatial, temporal or extreme event variability." The reviewer posits that it seems rather than addressing how to accurately estimate CDOM by modifying the algorithm, the authors highlight what is "abnormal" about these environments. The reviewer believes there is certainly room for this, but thinks the authors would be better served by focusing on how their algorithm could be adapted for these environments, rather than subjectively classifying environments where the algorithm does not perform well. The reviewer asserts the authors emphasize how the sensitivity of the instruments used detects these changes, but there is no analysis for how this increased sensitivity can be used to develop more capable algorithms.

Authors Response 38: Within the manuscript plus the Authors Responses herein, the Λ_{780}^{320} algorithm is unchanged. What is allowed to change is the optical complexity of the data used to evaluate the algorithm. Furthermore, many of the water masses that are part of the optical complexity in the subcategories are properly labeled considered abnormal, e.g., hypersaline, alkaline, or polluted lakes. They were sampled with the purpose of providing extreme data to quantify how the performance of the algorithm is degraded by such water masses. The results presented in Authors Response 37 for the universal, fifth, and sixth data sets establish that the algorithm is sufficiently robust to provide accurate results even in the presence of such water masses and confirm a more capable algorithm is not needed.

In regards to the sensitivity argument, the sensitivity of the C-OPS and C-PrOPS instrumentation becomes critical in providing confidence in the derived relationships. Using other instrumentation (e.g., the Satlantic HyperPro II), demonstrates this, in that the curated NOMAD dataset includes what the manuscript demonstrates to be aberrant measurements. This is not obvious when using instruments with lower sensitivity, because the variance is large enough that it is not clear whether these are true outliers. The authors think this point is adequately expressed in the manuscript without explicitly calling out the shortcomings of specific instruments, which would be necessary to adequately discuss the per-instrument sensitivity (performance) of different datasets. Consequently, the manuscript will not be modified further.

Reviewer 2 Comment 19: The reviewer cites P18 L27–28 regarding how local wind conditions could elevate the values associated with a typical bloom into atypical concentrations and asserts this calls into question the purpose of classifications. The reviewer also notes these are conditions that will be observed, either through *in situ* or satellite observations and wonders if the algorithm could be improved by factoring in wind conditions.

Authors Response 39: The authors believe this is a good example as to why the subjective classification process is powerful, because it includes an external important forcing mechanism. For example, if the data from the observations involved in the local wind conditions cited above are used to evaluate the performance of the OC3M6 algorithm, the HPLC TChl a concentration on the sheltered upwind side of a wind-blown polluted lake is $67.484\,\mathrm{mg\,m^{-3}}$ and $1{,}116.512\,\mathrm{mg\,m^{-3}}$ on the opposite downwind shore. The uncertainty in the OC3M6 algorithm for the lee side is 91.2% and for the opposite shore it is 1,555.0%.

The authors are unaware of any example where local wind conditions were used to improve chlorophyll retrievals by factoring in wind conditions, but if the data from the polluted wind-blown lake were used to validate a chlorophyll algorithm, it is anticipated that a notation regarding anomalous concentrations from wind effects would likely be appreciated. Without an effort associated with wind-blown corrections, the algorithm presented in the manuscript for the wind-blown polluted lake for $a_{\mathrm{CDOM}}(440)$ has an uncertainty on the lee side of the polluted lake of 24.5% and 36.4% on the opposite shore. These numbers are not so large as to render the $a_{\mathrm{CDOM}}(440)$ algorithm useless, but this does occur for the OC3M6 algorithm. Consequently, no modifications will be made to the manuscript.

Reviewer 2 Comment 20: The reviewer cites P20 L22-24 in regards the phenomenon that as end members are brought spectrally closer together, the range of expression available to distinguish two similar but optically different water masses decreases and suggests the mechanics of this could be explored and explained.

Authors Response 40: This phenomenon is easily understood by studying Fig. 8, wherein the range in the optical axis for the $K_d(313)/K_d(875)$ algorithm is greater than the range in the optical axis for the $K_d(412)/K_d(670)$ algorithm. The nuances of legacy algorithms is not a principal focus of the study and the material presented in the manuscript is deemed sufficient, so the manuscript is not modified.

Reviewer 2 Comment 21: The reviewer cites P21 L16-17 "Application of Λ_{670}^{412} data to the corresponding algorithm in Fig. 8 results in 13 observations with negative (predicted) $a_{\mathrm{CDOM}}(440)$ values, which are removed to leave 212 unique stations. This process demonstrates how end-member algorithms can be used to quality assure optical data in archives (Sect. 3.7)." The reviewer also cites P23 L2-4 "With respect to the algorithm, the increased bias, variance, and 13 negative derived values obtained with NOMAD data (which is a small, quality controlled subset of the larger NASA SeaBASS archive) in clearer waters suggests the legacy data are degraded by sampling artifacts." The reviewer then questions if this is an issue with the algorithm or the measurements.

Authors Response 41: The preponderance of evidence suggests that it is an issue with the legacy instrumentation. For the same geographical region under similar conditions, the C-OPS clear and turbid water data conform to the algorithm. The independent validation with the NOMAD dataset shows that the slope of the algorithm fit for the turbid partition of the NOMAD dataset is consistent with the slope found in the global algorithm perspective, but that the slope of the algorithm fit for the clear partition of the NOMAD dataset is significantly different (P22 L25–29).

Classification of the water mass can help assess whether the NOMAD issue arises from the algorithm or from the measurements, because the manuscript shows that algorithm performance varies with class assignment. Because the metadata for subjective classification does not exist for NOMAD, the objective classification scheme was applied to NOMAD and found the number of data in each class as follows: N_1 6, N_2 13, N_3 135, N_4 49, N_5 0, and 9 were unclassified. All of the turbid data were in classes N_3 and N_4, but the clear data were included classes N_1 and N_2 plus the 9 observations that were not classified. This means the slope of the clear partition was determined with 19 points that were classified and 9 that were not, which accounts for the poor performance with respect to the algorithm. It also suggests that the measurements were the issue with the NOMAD data, because the spectra could not be classified.

This example of using objective classification as an analytical or investigative tool will be briefly summarized and added to the manuscript at the very end of Sect. 3.7 (P27 L13).

Reviewer 2 Comment 22: The reviewer cites P23 L14-15, "The data set established herein has an extensive number of observations directly suitable for validation exercises (Figs. 3 and 9) plus 15 subcategories (Sect. 2.5) of potentially (but not automatically) problematic water bodies (Figs. 4–7), with the latter determined subjectively." The reviewer asserts the authors acknowledge that "problematic" water bodies were determined subjectively, and these observations do not agree well with the algorithm while questioning a) if these observations span natural environmental variability that can be observed, and b) outside of directly observing human structures (e.g., reflectance of a shipwreck visible from surface waters), why did the authors not consider how to retrieve valid $a_{\mathrm{CDOM}}(440)$ values for these waters, and rather chose to assume the algorithm works very well and these waters are problematic? The reviewer also notes it isn't clear how the algorithm can be used to determine whether legacy data are valid or not, because its performance is based on these subjective classifications.

Authors Response 42: The authors do not acknowledge that problematic waters were determined subjectively. As stated in the manuscript and by the reviewer, subcategories were considered to be potentially, but not automatically, problematic. Also, the authors do not acknowledge that the subcategory data do not agree well with the algorithm, because this phrase does not appear in the manuscript. Furthermore, and as presented in Sect. 4 (plus Authors Response 12 and 23), data from the subcategories were incrementally added to assess algorithm performance and all of the algorithm assessments do not involve any description wherein performance is described as being degraded by observations that do not agree well with the algorithm, including new assessments presented in Authors Response 37.

All optical profiling was to the depth of the 10% or 1% light level while remaining above the bottom depth, so no data were contaminated by bottom or manmade structures. All the optical data retrieved valid $a_{\mathrm{CDOM}}(440)$ values using standard processing, except the UV domain for White Lake required channel-by-channel processing and are considered estimated, which is why they were omitted from the fourth, but nonetheless comprehensive data set used as a comprehensive evaluation of the algorithm. The manuscript contains no statement that the algorithm works very well; all algorithm evaluations are provided in terms of quantified performance.

The ability of the algorithm to determine whether legacy data are valid was demonstrated in the manuscript with additional refinements in Authors Response 41. The argument that algorithm performance is based on subjective classifications requires the understanding that the non-conservative water masses are likely outside the range in the gradient of a constituent. As stated in the Introduction, the natural range of variability in water masses that may be exceeded by extreme events (but this does not imply that those events are not "natural," just that they are statistically anomalous). An objective classification scheme would still identify these data as "anomalous," but without ancillary (non-optical) data, it may not be obvious whether the anomalies are due to unusual water properties or lack of adherence to the protocols for measuring optical properties (see Authors Response 41).

The authors therefore suggest that if the optical data do not adhere to the algorithm, the first step would be to determine whether there are other factors (subjectively described within the 15 subcategories outlined in this manuscript). If there are no discernible reasons for the data to appear "anomalous" then it strongly suggests that there is an issue with data collection (see Authors Response 41), or that some environmental stressor not captured in the 15 subcategories resulted in an anomalous situation. In other words, the authors are not suggesting that strict adherence to the algorithm is a criteria by itself, but rather deviation from the algorithm should trigger additional scrutiny of the full dataset.

Clarification on these points will be improved through an addition to manuscript Sect. 4 (in P31 after L16) as described in Author Response 43.

Reviewer 2 Comment 23: The reviewer cites P24 L7-8, "Consequently, a subcategorization scheme based on the optical measurements alone might be advantageous to the validation process, particularly for archival data." This is assuming that CDOM should behave conservatively across water masses? It seems the very point the paper is making is that abnormal environments produce CDOM of a different spectral nature. This is important. Why have the authors not attempted to accurately estimate this variability?.

Authors Response 43: One point of the study is that for the vast majority of water bodies (by surface area), a single global algorithm effectively retrieves CDOM. For a subset of anomalous cases, there are two potential explanations for deviations (anomalies) from the expected fit: a) there is an unusual environmental factor occurring (and 15 subcategories are provided for evaluation, most of which do not significantly degrade the global algorithm), or b) the data were collected improperly (either the diffuse attenuation coefficients or the CDOM absorption coefficient values). Application of the algorithm to historical (archival) data will identify outliers, and this would help guide a more careful analysis of the preponderance of evidence for putting those data into one or the other category (i.e., the water mass is truly anomalous or the data collection is suspect).

To address the points in this comment, as well as for Reviewer 2 Comment 22, the manuscript will be modified in P31 after L16 by adding the following text: "Screening of newly collected or archival data with respect to a selected algorithm can be accomplished by initially flagging data points more than 12% from the expected relationship, and then more carefully examining those points using both objective and subjective criteria (based on available metadata) to determine whether the results are expected, or are more likely to indicate a problem with data collection procedures."

Reviewer 2 Comment 24: The reviewer cites P23 L18-24, "The author's reference clustering analysis, particularly fuzzy clustering presented by Moore et al. (various years)" and questions why subjective categorizations were used rather than an objective approach such as that of Moore et al.? The reviewer also suggests it would also be useful to reference this work earlier, perhaps in the introduction, and discussing the need for classifications to build effective algorithms could also be elaborated.

Authors Response 44: To provide insights into the dynamics of the observed K_d spectra, all data were first categorized into subcategories based on a subjective classification but using sampling information necessary to better understand optical properties, which would otherwise not be known. The robustness of the subcategories are logically supported by an objective classification using fuzzy c-means (FCM). Please note that FCM is different from a crisp or hard classification such as k-means. The authors believe that these two steps are required to appropriately interpret the results. Following the reviewer, a brief description about necessity of both subjective and objective classifications was added to Sect. 1 (P4 L18–25).

Reviewer 2 Comment 25: The reviewer cites P26 L11-14, "The decrease in the percent composition of the validation quality data as a function of increasing class number (N_1–N_5) is an indicator of the difficulty of validating an algorithm within increasingly complex waters. The recurring contribution of a relatively small number of principal subjective subcategories to the gradient in optical complexity starting with N_2 and then continuing for N_3–N_5 confirms the original subcategory approach has merit." The reviewer then posits that conversely, the end member approach only performs well in bodies of water with little optical variability. The reviewer states it isn't clear why a subjective deconstruction of water bodies was used versus a fuzzy clustering approach where alternate relationships between K_d and $a_{\mathrm{CDOM}}(440)$ were explored, noting that fuzzy clustering approaches use an objective classification scheme to separate out waters with the intention of providing a framework where different algorithms can be applied, and questions why was that not explored.

Authors Response 45: One of the objectives of the study was to examine a global algorithm for estimating $a_{\mathrm{CDOM}}(440)$ that works for a diversity of water masses, which is a different approach compared to what the reviewer mentioned, e.g., Hieronymi et al. (2017). While the authors explained "anomalies" using both subjective and objective classifications, overall performance of the algorithm was always superior to or in keeping with global algorithms in use by the community of practice—even when the present study incrementally added in increasingly complex water masses beyond the capabilities of existing global algorithms (see Authors Response 12, 23, and 37). Consequently, no additional modifications were made to the manuscript.

Reviewer 2 Comment 26: The reviewer cites P26 L28 and the use of "evolving conservatively" and then questions a) if this is being used to represent anything that is a natural process within the water column, with no added optical constituents, and b) whether photo- and microbial degradation of CDOM considered a conservative process. The reviewer suggests a clearer discussion of Case 1 and Case 2 waters and how that classification relates to the classification used here would clarify this.

Authors Response 46: The manuscript already answers the two questions on P3 L28 to P4 L3 wherein evolution that is within a constrained (natural) range of a water mass are considered conservative. The requested clarifications regarding case-1 and case-2 waters already appear in the manuscript on P25 L18 to P26 L1, but additional small modifications will be added to P26 L2–10 to improve clarity and provide completeness.

Editor 1 Comment 1: The editor suggests a more complete characterization of the dataset by discussing the range in the CDOM absorption spectral slopes and/or slope ratios that are relevant to the quality and composition, rather than amount, of CDOM.

Authors Response 47: The authors agree, and for all CDOM data, S for the wavelength range 350–500 nm was recalculated by fitting a nonlinear least-squares model to $a_{\mathrm{CDOM}}(\lambda)$. The choice of the spectral range was based on the fact that all data within the spectral domain have sufficient SNR, but no saturation issue (calculation of a spectral slope in a shorter spectral domain, e.g., 275–295 nm for all data was not possible due to saturation when using a long path length). This is a compromise, because long path-length data provide more accurate results compared to those obtained with a short path length (e.g., 10 cm) in the visible spectral domain. The new analysis regarding spectral slopes showed the range was comparable to similar global analyses spanning the majority of marine waters, e.g., Aurin et al. (2018) and Grunert et al. (2018). In addition, the application of algorithm performance presented Figs. 3–8 was not improved using spectral slope data as an alternative to $a_{\mathrm{CDOM}}(440)$, i.e., the deviations of data points using $S(350--500)$ was not better explained. Consequently, modifications to the revised manuscript were restricted to the range discussion with appropriate comparisons to the literature. The necessary text was added to Sect. 3 (P14 L15–24 in the revised manuscript).

Editor 1 Comment 2: The editor indicates the the terms “atypical” or “anomalous” were a source of confusion and suggests a revision throughout the manuscript to clarify that the 15 data subcategories mentioned in the paper are categories of “more complex” water masses that deviate from the predicted aCDOM vs Kd-ratio relationship.

Authors Response 48: The authors agree and have removed the identified vocabulary and made the necessary small text changes throughout the revised manuscript.

Editor 1 Comment 3: The editor indicates some additional information is needed in the Methods section to assess/quantify the impact of freezing/thawing on CDOM optical properties.

Authors Response 49: The authors agree and have modified Sect. 2.4 to include new material to address this point (P8 L1–3 in the revised manuscript).

Editor 1 Comment 4: The editor encourages the authors to reduce the length of the manuscript.

Authors Response 50: The authors agree and revised the manuscript to significantly shorten it. Unfortunately, the reviewer and editor comments also requested that a substantial amount of new material be added. Despite the requirements for new material, the main body of the revised manuscript is nonetheless a bit more than two pages less than the original.

The marked up version of the resubmitted manuscript is provided below, with all changes to the original document shown in red text.

Respectfully yours,

Stanford Hooker

A Global End-Member Approach to Derive $a_{\mathrm{CDOM}}(440)$ from Near-Surface Optical Measurements

Stanford B. Hooker†
NASA Goddard Space Flight Center
Ocean Ecology Laboratory/Code 616.2
Greenbelt, Maryland 20771

Atsushi Matsuoka
Université Laval
Takuvik Joint International Laboratory
Québec City, Canada G1V 0A6

Raphael M. Kudela
University of California Santa Cruz
Ocean Sciences Department
Santa Cruz, California 95064

Youhei Yamashita
Faculty of Environmental Earth Science
Hokkaido University
Sapporo, Japan 060-0810

Koji Suzuki§
Faculty of Environmental Earth Science
Hokkaido University
Sapporo, Japan 060-0810

Henry F. Houskeeper
University of California Santa Cruz
Ocean Sciences Department
Santa Cruz, California 95064

† *First Corresponding Author:* `stanford.b.hooker@nasa.gov`

§ *Second Corresponding Author:* `kojis@ees.hokudai.ac.jp`

Declaration of conflict of interest: None

November 18, 2019 Revised and resubmitted to Biogeosciences after review.

Abstract

This study establishes an optical inversion scheme for deriving the absorption coefficient of colored (or chromophoric, depending on the literature) dissolved organic material (CDOM) at the 440 nm wavelength, which can be applied to global water masses with near-equal efficacy. The approach uses a ratio of diffuse attenuation coefficient spectral end members, i.e., a short and long wavelength pair. The global perspective is established by sampling "extremely" clear water plus a generalized extent in turbidity and optical properties that each span three decades of dynamic range. A unique data set was collected in oceanic, coastal, and inland waters (as shallow as 0.6 m) from the North Pacific Ocean, the Arctic Ocean, Hawaii, Japan, Puerto Rico, and the western coast of the United States. The data were partitioned using subjective categorizations to define a validation quality subset of conservative water masses, i.e., the inflow and outflow of properties constrain the range in the gradient of a constituent, plus 15 subcategories of more complex water masses that were not necessarily evolving conservatively. The dependence on optical complexity was confirmed with an objective methodology based on a cluster analysis technique. The latter defined five distinct classes with validation quality data present in all classes, but which also decreased in percent composition as a function of increasing class number and optical complexity. Four algorithms based on different validation quality end members were validated with accuracies of 1.2–6.2%, wherein pairs with the largest spectral span were most accurate. Although algorithm accuracy decreased with the inclusion of more subcategories containing non-conservative water masses, changes to the algorithm fit were small when a preponderance of subcategories were included. The high accuracy for all end-member algorithms was the result of data acquisition and data processing improvements, e.g., increased vertical sampling resolution to less than 1 mm (with pressure transducer precision of 0.03–0.08 mm) and a boundary constraint to mitigate wave focusing effects, respectively. An independent evaluation with a historical database confirmed the consistency of the algorithmic approach and its application to quality assurance, e.g., to flag data outside expected ranges, identify suspect spectra, and objectively determine the in-water extrapolation interval by converging agreement for all applicable end-member algorithms. The legacy data exhibit degraded performance (as 44% uncertainty) due to a lack of high-quality near-surface observations, especially for clear waters wherein wave-focusing effects are problematic. The novel optical approach allows the *in situ* estimation of an in-water constituent in keeping with the accuracy obtained in the laboratory.

Keywords: CDOM; absorption coefficient; inversion algorithm; oceanic; coastal; inland; global

1. Introduction

The colored (or chromophoric, depending on the literature) dissolved organic matter (CDOM) spectral absorption coefficient, $a_{\mathrm{CDOM}}(\lambda)$, where λ is wavelength, is widely used to investigate terrestrial and oceanic biogeochemical processes, as summarized in the review by Nelson and Siegel (2013). The selection of $a_{\mathrm{CDOM}}(440)$ as the parameter of interest is a consequence of the relationships between CDOM and the solar illumination of aquatic ecosystems, as follows: a) CDOM protects microorganisms from harmful ultraviolet (UV) radiation, albeit while reducing photosynthetically available radiation (Nelson and Siegel 2013); b) CDOM affects the heat content of a water mass, e.g., causing stratification for brown lakes (Houser 2006); c) CDOM supplies inorganic nutrients, i.e., ammonium, (Bushaw et al. 1996) and can be a source of labile organic substances (Mopper et al. 1991) through photochemical degradation and mineralization processes; and d) CDOM is a potentially useful parameter to distinguish and trace water masses in the coastal zone and open ocean (Nelson et al. 2007 and Tanaka et al. 2016).

This study evaluates whether a proposed algorithm (Hooker et al. 2013) for deriving $a_{\mathrm{CDOM}}(440)$ from a ratio of diffuse attenuation coefficient spectral end members, $K_d(\lambda_1)/K_d(\lambda_2)$, with the shortest wavelength denoted λ_1 and the longest denoted λ_2, can be applied to *global* water masses with equal efficacy. Typically, $a_{\mathrm{CDOM}}(440)$ is determined in the laboratory using an optical instrument and a water sample obtained *in situ*. In this study, the water sample is collected in temporal close proximity to in-water optical sampling used to derive $K_d(\lambda)$ from vertical profiles starting sufficiently close to the water surface to accurately derive the widest range of K_d wavelengths.

The Hooker et al. (2013) $a_{\mathrm{CDOM}}(440)$ algorithm is based on a straightforward principal, as follows: if a water mass is studied optically in a homogeneous near-surface interval of the water column, optical data products can be derived for all wavelengths and the most sensitive parts are the spectral end members. The end members exhibit the greatest range in values as a function of the absorption and scattering processes responsible for the attenuation of light and can be inverted to derive typical constituents as a function of changes in attenuation properties.

Although an ability to derive an in-water constituent from optical measurements provides a follow-on opportunity for airborne or satellite synoptic surveys, this is not a principal objective of this study. The reason for de-emphasizing remote sensing is the remote sensing instruments typically available do not provide the spectral range used herein, although a *legacy* pair of wavelengths are considered below. In addition, the principal parameter used here is not measured directly by a remote sensor. Although $K_d(\lambda)$ can be derived from remote sensing data for part of the needed spectrum (Cao et al. 2014), the inversion is incomplete and the introduced inaccuracies compromise a principal objective, which is to determine $a_{\mathrm{CDOM}}(\lambda)$ *in situ* with an accuracy commensurate with laboratory analyses. Finally, despite the success with high-spatial-resolution remote sensing platforms for studying coastal and inland waters (Palmer et al. 2015 and Mouw et al. 2015), this study includes water bodies too small to be studied by such platforms, and the reduced spectral resolution or range, coupled with the methods for characterizing $a_{\mathrm{CDOM}}(\lambda)$, have proved inadequate even within relatively large, lacustrine water bodies (Kutser et al. 2015).

The *global* perspective refers to a generalized concept of sampling a multitude of geographical areas and watersheds wherein three broad categories are sampled: open ocean, coastal zone (e.g., shelf waters, bays, estuaries, lagoons, etc.), and inland water bodies (e.g., rivers, lakes, reservoirs, wetlands and marshes, etc.). The near-surface viewpoint is not driven exclusively by the desire to produce data products at all wavelengths. The other reasons for sampling and deriving data products close to the water surface are as follows: a) establish a technique that can ultimately support remote sensing objectives as the technologies advance, wherein the spaceborne and airborne approaches obtain data products directly from the sea surface signal; b) use the same protocols for sampling and deriving data products for all water masses, so the widest dynamic range in water properties can be considered (the shallowest water depth sampled was 0.6 m); and c) improve the use of the global solar irradiance observations (obtained with a separate solar reference) in setting a constraint for the fitting procedures used to derive the in-water data products (Antoine et al. 2013), the effectiveness of which is related to how close the extrapolation interval is to the surface.

Using a homogeneous near-surface interval to derive all data products ensures the spectral interrelationships coincide with the same water used to determine the in-water constituents by laboratory analysis. The perspectives of natural changes and typical properties are also important, because some water bodies are not automatically assumed to have typical water properties. For example, endorheic lakes are enclosed, so ground seepage and evaporation are the principal outflow mechanisms with evaporation continuously concentrating constituents (Yapiyev et al. 2017). Over time, a narrowly defined ecosystem evolves to withstand the increasingly extreme conditions, and in some cases, higher-order life ceases to exist, e.g., the Dead Sea (Bodaker et al. 2010).

Endorheic lakes are an end point in the expression of water masses, because the range in the temporal gradient of a constituent, e.g., salt, is somewhat unbounded and the water body does not evolve conservatively due to the significant outflow versus inflow imbalance. For purposes exclusive to this discussion, a conservative water body is defined to have an inflow and outflow of properties that constrain the range in the gradient of a constituent. This natural range is usually established by seasonal factors, although unnatural or atypical stressors can add optical complexities, which may or may not be seasonal. Examples of the latter are anthropogenic sources (e.g., pollution or agricultural water diversion) and severe weather (e.g., typhoon-induced bottom resuspension in coastal ecosystems).

Consequently, other water bodies subjected to an unexpected stressor that allows an unbounded gradient in a constituent, e.g, long-term drought, are anticipated to not evolve conservatively and the constituents might be expressed as extreme values as a function of time. Once the stressor is no longer applied, the water mass evolves semi-conservatively, wherein the atypical properties are diluted or removed, and at some point in time the water body reverts to a conservative evolution, i.e., the gradient in the constituent is within an expected or natural range.

A global perspective is constructed with overlapping ranges in the natural gradient of the constituent and the optical inversion parameters within conservative water masses (Lee and Hu 2006). If the assembled dynamic range extends across a sufficiently dense sampling of clear to turbid water masses, an explicit sampling of every possible global water mass is not deemed necessary. The turbid-water endpoint of the dynamic range is somewhat undefined, because of present limitations in obtaining in-water optical measurements in extremely shallow or turbid waters (a case in point, White Lake, is presented below), but the clear-water endpoint is defined by the pure-water limit, i.e., no constituents (Morel 1974). Consequently, the dynamic range herein can only be extended in one direction and any turbid additions involve necessarily small volumes, so the global perspective is at most only marginally incomplete.

Based on the degree of complexity for a water mass not evolving conservatively, it is anticipated that such water masses are unsuitable for validating a global algorithm established to invert the optical properties of conservative waters. Whether a validation site is inappropriate is a function of the severity of the stressor creating the non-conservative evolution. For example, a short-term water diversion from a lake is expected to create a short-term complexity, whereas a long-term drought likely creates a time series of increasingly extreme values (Vazquez et al. 2011 and Guarch-Ribot and Butturini 2016). Consequently, the sampling objective used here was to obtain measurements in conservative water masses plus water bodies subjected to one or more stressors. To assess performance, both subjective and objective classifications of water mass complexity are included in the algorithm evaluation process.

2. Methods

The Hooker et al. (2013) study excluded lacustrine water bodies; the largest inland water masses were tidal estuaries. Consequently, the new validation data set includes a large variety of lakes and reservoirs, wherein some were selected precisely because compliance with the original (Hooker et al. 2013) algorithm was not anticipated. These nonconservative water bodies provide an important test of the algorithmic approach, because if they do not appear as outliers with respect to the original algorithm, the principles behind the algorithm are challenged.

To improve the quality of optical measurements obtained in near-surface waters, which is essential for studying shallow ecosystems, methodological advancements were included for this study. Consequently, the methods described herein are distinguished with respect to the original Hooker et al. (2013) research as follows: a) a significantly enlarged study area with new water body types (e.g., lakes and reservoirs, more numerous rivers, the marginal ice zone, etc.); and b) the use of more advanced optical technologies to improve sampling efficiency and data quality.

2.1 Optical Instrumentation

The optical instrument suite deployed for this study is a handheld, free-falling Compact-Optical Profiling System (C-OPS), as first described by Morrow et al. (2010), that measures the downward irradiance and upwelled radiance, $E_d(z,\lambda)$ and $L_u(z,\lambda)$, respectively, where z is depth. An above-water reference, sited to avoid shadows and reflections, simultaneously measures the global solar irradiance, $E_d(0^+,\lambda)$, where 0^+ indicates above the water surface. This configuration was deployed by Hooker et al. (2013), except the study documented herein used advanced radiometers with three gain stages rather than two. The majority of profiles were obtained with the Compact-Propulsion Option for Profiling Systems (C-PrOPS) plus a conductivity sensor for improved water mass characterization. The former uses two small digital thrusters to maneuver the backplane (Hooker et al. 2018a) beyond the influence of platform perturbations, which does not remove the self-shading effect, thus, necessitating an $L_u(\lambda)$ correction (Gordon and Ding 1992). Hooker (2014) provides the negative consequences of common positioning alternatives.

A transparent drawing of the next-generation C-OPS with C-PrOPS is presented in Fig. 1. When the weak thrust holding the (slightly) negatively buoyant profiler at the surface is removed, one or more bladders in the hydrobaric buoyancy chamber slowly compress and increase the near-surface loitering of the profiler, which results in a vertical sampling resolution (VSR) to within 1 cm or less. The VSR is defined as the vertical extent of the extrapolation interval used to derive the data products, e.g., $K_d(\lambda)$, divided by the number of retained data points in the interval, wherein retention requires planar orientation to within 5° of vertical. For coastal and inland waters, the average VSR was 6.0 mm, but for very shallow or turbid waters, the average VSR was 0.9 mm. In comparison, the Hooker et al. (2013) study had a VSR of approximately 10.0 mm. For the open ocean, the average VSR was 12.9 mm, because open-ocean profiles were in a more turbulent wave field, so the profiler was ballasted to sink faster and descend deeper. The open ocean deep mixed layers mean a slightly coarser vertical resolution is not a limitation.

The two digital thrusters have the same cant angle with respect to the vertical, which directs the weak turbulence from the thrusters downward and below the irradiance instrument, thereby ensuring both light apertures are observing undisturbed water; the opposite occurs if thrust is reversed. To steer the backplane like a remotely operated vehicle, differential thrust is applied to the two thrusters (Hooker et al. 2018a) and allows for real-time positioning adjustments, which is a significant advantage in shallow waters, e.g., away from a shoreline or within a wetland.

Once the profiler reaches the desired position for obtaining a vertical profile of measurements, or *cast*, it is kept in position by maintaining weak forward thrust. While at the surface, the pressure transducer measures atmospheric pressure right before a profile commences, which allows a pressure tare for every cast and improves the accuracy of depth measurements (Hooker 2014). The pressure transducers used herein have a depth resolution, in terms of precision, of 0.03–0.08 mm in all water masses. When thrust is removed, the thruster-induced bias in the roll axis relaxes (the pitch angle is negligible, because prior thrust aligns the backplane with almost no pitch angle), and the profiler descends with stable tilts (Hooker et al. 2018a). Unlike rocket-shaped profilers (Hooker et al. 2001), the profiler has no significant righting moment and the planar orientation of the radiometers is maintained from the start of data acquisition, which significantly improves the VSR. In deep waters, the thrusters are used at the bottom of the cast to reduce the time between casts; in very shallow waters, the profiler was hauled closer to the surface before the thrusters were used to prevent resuspension of bottom material. All profiles were obtained in waters wherein the 10% light level was above the bottom depth to ensure all data products were uncontaminated by bottom reflections.

Seven different instrument suites including a next-generation *hybridspectral* profiler with fixed-wavelength and hyperspectral detector components plus a C-PrOPS (Hooker et al. 2018b) were used for this study. The fixed-wavelengths of the radiometers had similar configurations such that all measured the same nine spectral end members from 320–412 nm and 670–780 nm, plus six common wavelengths (Table 1). All optical instruments were calibrated at the same manufacturer facility with traceability to the National Institute of Standards and Technology (NIST) as described by Hooker et al. (2018b). NIST traceability is a requirement of the NASA Ocean Optics Protocols (hereafter, the Protocols). The Protocols set the standards for calibration and validation activities (Mueller and Austin 1992), which were revised (Mueller and Austin 1995) and updated over time (Mueller 2000, 2002, and 2003).

A comparison of C-OPS acquisition with and without thrusters (Hooker et al. 2018a) verified the former improved efficiency in all waters by a factor of two or more, plus either no or minor adjustments to the extrapolation interval used to derive data products for replicate casts were needed (thrusters minimize the negative influences of heterogeneity across all wavelengths). The improved efficiency yields a closer temporal matchup between the collection of optical profiles and the water sample. Approximately 98% of water samples were obtained from the surface using a bucket. For some inland waters, the profiler was launched from a shoreline or dock e.g., when the boat ramp was out of service (due to drought, flooding, invasive species regulations, etc.). If a water sample could not be otherwise retrieved from the profiling location, the Compact-Profiler Underway Measurement Pumping System (C-PUMPS) was used (Hooker et al. 2018a). C-PUMPS provides a $20\,\mathrm{ml\,s^{-1}}$ flow rate from the profiler and fills a 1 l container in less than 1 min.

2.2 Field Sampling

The Hooker et al. (2013) study area was the Beaufort Sea in proximity to the Mackenzie River outflow, the Gulf of Maine and vicinity, including major portions of its inland watershed plus minor watershed drainage from smaller rivers and a saltwater marsh. A validation data set from observations made in U.S. coastal waters within the southern Mid-Atlantic Bight were also used. Neither of the two data sets included typical lacustrine water masses. The new sampling area for the study herein included the western U.S. (i.e., California, Oregon, Washington, Nevada, Utah, and Idaho), Hawaii, Puerto Rico, Japan, the western North Pacific Ocean (e.g., the Kuroshio and Oyashio Currents), the central North Pacific Ocean, the Bering Sea, the Chukchi Sea, and the Beaufort Sea (Fig. 2). The latter is the only region that slightly overlaps the Hooker et al. (2013) study. The new data set includes sampling in a wide diversity of inland rivers, lakes, and reservoirs, including hypersaline and alkaline lakes.

The new field data are divided into the aforementioned three primary categories according to whether or not the sampling station was in the open ocean, coastal zone, or inland waters. The open ocean is defined as offshore waters with a water depth exceeding 200 m. The coastal zone includes near-shore bathymetry of 200 m or less, wherein the adjacent saline waters and shorelands strongly influence each other, and includes islands, bays, deltas, transitional and inter-tidal areas, salt marshes, wetlands, beaches, etc. Inland waters are all other water bodies landward of the coastal low-water line, which are predominantly—but not exclusively—fresh lacustrine and riverine ecosystems.

Twenty-five campaigns spanning 29 April 2013 to 25 January 2017 were conducted with 318 stations occupied and 1,230 vertical profiles obtained, which were executed as a minimum of three sequential casts at each station. A majority of the optical data (733 casts) were obtained with C-PrOPS (Table 1) in all three primary categories, whereas optical sampling without thrusters (497 casts) was almost exclusively in the open ocean and coastal waters with minimal heterogeneity and deep mixed layers. Duplicate, and sometimes triplicate, water samples were collected at each C-PrOPS station. For open-ocean campaigns in the Pacific Ocean and Arctic, which included some coastal waters, a single seawater sample was usually collected. A selected volume of each water sample was filtered through a 0.22 µm filter under a gentle vacuum and collected in appropriate (e.g., pre-combusted) clean glass vials or bottles. Typically, CDOM absorption was measured within a few hours after sampling. In some cases when this was not possible, samples were stored at $-20\,^{\circ}$C or less until subsequent laboratory analysis at a shore facility (Sect. 2.4). While Fellman et al. (2008) reported that freezing changed the chemical composition of DOC for freshwater (stream) samples, Hancke et al. (2014) found no such effect for Arctic marine samples, nor was any systematic bias observed between the subset of frozen samples compared to the full suite of samples used for this analysis.

The surface water sample was obtained as quickly as possible after three optical casts were performed. In some cases when the heterogeneity or turbulence of the water mass was considered to be excessive, an additional three optical casts were executed immediately after the water sample was collected. The determination of excessive conditions was based on the stability achieved in optical variables (e.g., the average vertical tilt in the upper 2 m of the water column, changes in the 10% light level depth, etc.) during data acquisition for the first three casts.

2.3 Optical Data Processing

All optical data products discussed herein, e.g., $K_d(\lambda)$, were estimated in a near-surface interval of the water column with homogeneous properties as confirmed with physical data and analysis of the linearity of extrapolations provided by the processing scheme. The processor used here is based on a well-established methodology (Smith and Baker 1984) that Hooker et al. (2001) showed is capable of agreement at the 1% level within an international round robin, when the processing options are as similar as possible and both data acquisition and processing strictly adhere to the Protocols. Summary details of the data acquisition and processing capabilities are provided in Antoine et al. (2013), Hooker (2014), and Hooker et al. (2018a, 2018b, and 2018c), so only brief overviews are presented herein.

In-water radiometric parameters in physical units are normalized with respect to separate, but simultaneous, $E_d(0^+, \lambda, t)$ measurements, with t expressing time dependence. After solar normalization and $\pm 5°$ tilt filtering, a near-surface homogeneous portion of $E_d(z, \lambda)$ centered at depth z_0 and extending from $z_1 = z_0 + \Delta z$ and $z_2 = z_0 - \Delta z$ is established separately for the blue-green and red wavelengths; the UV and near infrared (NIR) wavelengths are included in the blue-green and red intervals, respectively. Both intervals begin at the same shallowest depth, but the blue-green interval is allowed to extend deeper if the extrapolation linearity, as determined statistically, is thereby improved (this only occurs in oligotrophic, optically simple, waters with deep mixed layers). The negative value of the regression slope yields $K_d(\lambda)$, which is used to extrapolate the fitted portion of the E_d profile to $z = 0^-$.

A principal benefit of profile data with a high VSR is that the aliasing caused by wave-focusing effects (Zaneveld et al. 2001) can be significantly reduced during data processing. The separately obtained above- and in-water E_d values at $z = 0^+$ and $z = 0^-$, respectively, can be compared using

$$E_d(0^-, \lambda) \;=\; 0.97\, E_d(0^+, \lambda), \tag{1}$$

where the constant 0.97 represents the applicable air-sea transmittance, Fresnel reflectances, and the irradiance reflectance (Morel et al. 2007). The distribution of light measurements at depth z influenced by wave-focusing effects do not follow a Gaussian distribution, especially during clear-sky conditions, wherein the amplitude of the brightened signals exceed the companion darkened signals. Consequently, arithmetic averaging is not appropriate and linear fitting of E_d in a near-surface layer is poorly constrained, especially if the number of samples is small.

The appropriateness of the E_d extrapolation interval, initially established by z_1 and z_2, is evaluated by determining if (1) is satisfied to within the 2.3–2.7% uncertainty ($k = 2$ coverage factor) of the optical calibrations (Hooker et al. 2018b); if not, z_1 and z_2 are redetermined—while keeping the selected depths within the shallowest homogeneous layer possible—until the disagreement is minimized (usually to within 5% to include some inevitable variance from natural processes to the calibration uncertainty). In this procedure, selection of the near-surface extrapolation interval uses a boundary condition or constraint (Antoine et al. 2013), wherein the central tendency of the distribution of data within the extrapolation interval, which are typically subjected to wave-focusing effects, satisfies (1).

The linear decay of all light parameters in the selected near-surface layer are then evaluated, and if linearity is acceptable, the entire process is repeated on a cast-by-cast basis. Subsurface quantities at null depth are obtained from the slope and intercept given by the least-squares linear regression versus z within the extrapolation interval specified by z_1 and z_2. A secondary benefit of profile data with a high VSR is that the extrapolation interval can have a restricted vertical extent, but still have sufficient data to satisfy (1) and produce data products at all wavelengths. This is an important advantage in optically complex water masses, which are usually turbid and shallow.

2.4 Water Sample Processing

For the Pacific Ocean samples and approximately half of the Arctic samples, the absorption spectrum of CDOM was determined using a spectrophotometer (Shimadzu UV-1800) according to Yamashita et al. (2013). Briefly, after the water sample was thawed and reached room temperature, the spectral absorbance was measured from 200–800 nm at 0.5 nm intervals with a 10 cm quartz-windowed cell. Absorbance spectra of a blank (Milli-Q water) and samples were obtained against air, and a blank spectrum was subtracted from each sample spectrum. The blank-corrected absorbance spectrum was baseline-corrected by subtracting average values ranging from 590–600 nm (Yamashita and Tanoue 2009), and then converted to the absorption coefficient (Green and Blough 1994). A single absorbance analysis was generally carried out for the open ocean samples, with an average accuracy from replicates of 2.5%.

For the other Arctic samples, CDOM absorption coefficients were determined using an UltraPath liquid waveguide (Matsuoka et al. 2012 and 2017). Briefly, CDOM absorbance for a filtrate (less than 0.2 μm) was measured relative to a reference water within a few hours after sampling. The reference water was prepared in advance using pre-combusted pure salt (450 °C for 4 h) with Milli-Q water to adjust salinity within ±2 between a sample and a reference to correct for the refractive index effect. A 2 m optical path was used for all waters except some coastal sites wherein a 0.1 m path was used (Matsuoka et al. 2012 and 2014). After baseline correction, absorbance was converted into absorption coefficients by including the optical path length. The detection limit was within $0.001\,\mathrm{m}^{-1}$ (Matsuoka et al. 2017).

For the western U.S. coastal and inland waters, water samples were passed through a $0.2\,\mu\mathrm{m}$ syringe filter (Whatman GD/X) and absorbance of CDOM was measured on either a Cary Varian 50 spectrophotometer using a 10 cm quartz cell or an UltraPath liquid waveguide spectrometer with 2 m path length. The syringe filter was rinsed with sample prior to collection, with the sample stored in an amber, acid-washed and combusted (450 °C for 4 h) glass vial with Teflon septa, and kept in the dark at 4 °C until analysis. Absorption spectra of the filtered samples were measured using ultrapure water from a Millipore Milli-Q A10 pure water system with UV to reduce total organic carbon to less than 10 ppb. The absorption coefficient (calculated as absorbance divided by path length, multiplied by 2.303 to convert to natural log units) at 440 nm to represent CDOM abundance was estimated using the Single Exponential Model (SEM) for absorption from 300–700 nm as described by Twardowski et al. (2004). The three different methods used in this study for determining CDOM absorption do not influence the results (see a sensitivity analysis in Sect. 4). For all the CDOM data—regardless of the collection and processing method—the spectral slope for the wavelength range 350–500 nm was calculated by fitting a nonlinear least-squares model to $a_{\mathrm{CDOM}}(\lambda)$.

2.5 Data Subcategories

Algorithm validation requires an assessment prior to data collection whether or not the sampling violates the assumptions made to create the algorithm. For this study, water masses wherein the evolution of $a_{\mathrm{CDOM}}(440)$ or $K_d(\lambda)$ was considered conservative, i.e., within the likely range in the gradient of such properties because no stressors to challenge that perspective were evident, are considered validation quality with primary categories of open ocean, coastal zone, or inland waters. Those water masses thought to contradict the algorithmic approach, because one or more stressors challenging the conservative evolution perspective were evident prior to sampling, are subcategorized to exclusively assess algorithm performance as more complex water masses are included in validation, as follows:

1. Waters closer to an ice field contain meltwater properties, which freshen the neighboring water body and can result in additional particles or compounds not usually found in the parent water mass.
2. Waters farther from an ice field, but within proximity, containing lesser amounts of meltwater properties.
3. Resuspension occurs naturally when a sufficient flow (e.g., an ebb or flood tide) or turbulent wave field (e.g., created by sufficiently strong winds) interacts with shallow bottom sediment to create concentrations of constituents that would otherwise not be present; it occurs unnaturally when a boat propeller (or other mechanical device) churns up shallow bottom sediment (e.g., in a harbor, marina, or navigation channel).

4. A refilled lake experiences a rapid inflow of alluvium (e.g., gravel, sand, silt, and clay) from riverbeds and eroding banks, plus floating and partially submerged debris, that can also resuspend bottom sediment. If the refilled lake is a controlled reservoir and exceeds the normally maintained fill level, new lake bottom is added, which can be a source for additional, perhaps atypical, water constituents in terms of type or concentration.
5. A drought-stricken lake has a longer residence time (the amount of time for the time-elapsed outflow to equal the lake volume) than normal, because once the water level remains below the overflow elevation, evaporation and ground seepage are the primary outflows. Increased residence time can concentrate constituents, plus dried and exposed bottom material can be resuspended into the shrinking lake volume due to wind or rain.
6. A harbor (or marina) is a docking facility, usually in shallow water, for vessels of varying sizes. Such facilities can be a source of pollutants and bottom resuspension, and typically include structures (breakwaters, jetties, piers, etc.) for shelter from severe weather, which can alter residence times by restricting water exchange.
7. A harmful algal bloom (HAB) is a toxic or noxious algae in a concentration producing a deleterious effect on humans or the environment. HABs are usually influenced by chemical, physical, and biological factors.
8. A wetland (plus marsh or mangrove) filters dissolved and suspended water constituents (e.g., from tidal cycles, weather events, etc.) through settling and plant consumption, but might not completely remove them.
9. A polluted water mass is contaminated from an anthropogenic source that alters the natural water properties.
10. An alkaline (or soda) lake has limited biodiversity due to an elevated pH of 9–12 with high carbonate and complex salt concentrations affecting the solubility and toxicity of chemicals and heavy metals (Grant 2006).
11. A hypersaline lake contains high concentrations of sodium chloride (or other salts) surpassing seawater, which limits biodiversity to organisms tolerating high saline levels (e.g., Mono Lake had a salinity of about 50).
12. A river mouth is where significant amounts of alluvium are deposited into a larger water body (e.g., a delta).
13. An atypical bloom is a generic case of high biomass based on local reports evaluated with respect to typical temporal and spatial conditions, which may involve weather effects (e.g., wind) concentrating algae advectively.
14. An invasive species is an introduced plant, fungus, or animal that is not native to a water body and is anticipated to alter the heretofore established properties and perhaps with a significantly negative outcome (e.g., damage to the environment, economy, or health of organisms, including humans).
15. A parent water mass modifier is a localized alteration of water properties, e.g., a creek inflow into a lake, and demonstrates the sensitivity of the methods used herein to distinguish small changes.

The above 15 subcategories plus the original validation quality category results in 16 categorizations, and as the former more complex waters are incrementally added to the latter, quantitative cause-and-effect validation scenarios are created (Sect. 4). If a sample was applicable to more than one subcategory, e.g., a wetland can experience resuspension from tidal currents, a dominant subcategory was selected based on observations prior to sampling. Categorization ambiguities are not worrisome, because complex water masses are a small fraction of global ecosystems.

The proximity to ice (farther and closer) subcategories are based on the relative position of safely operating a small vessel in and around an ice field. Sampling was usually as close to the ice as possible, and then as far from the ice within line of sight of the larger ship the small boat was launched from. In comparison, categorizing a refilled lake is a straightforward comparison of the water level datum available from local authorities with respect to the outflow elevation and historical norms. All refilled lakes were at 100% capacity or more, e.g., Washoe Lake was overfilled.

The categorization of bottom resuspension is primarily based on visual evidence, wherein resuspended particles are visible and produce a significant change in water color (e.g., Akkeshi Bay after the passage of typhoon Vongfong). Although a subset of sampling obtained in harbors could be classified as resuspension stations, a harbor (or marina) is identified based on local identification of such facilities. Similarly, wetlands (plus marshes and mangroves) are identified based on navigation charts, maps, and local descriptions. Hypersaline (endorheic) lakes are similarly categorized by state and local authorities (e.g., Mono Lake, Great Salt Lake, and Salton Sea), as are alkaline lakes (e.g., Mono Lake, with dual classification, plus Borax Lake and Soda Lake).

Categorizing drought-stricken lakes relies on local authorities reporting lake elevations and inflow water volumes with respect to historical norms. Examples from 2015 are as follows: a) Shasta Lake water storage was 56% below normal storage; b) Lake Almanor was 118.5 ft (36.1 m) below normal elevation; c) the Truckee River flow into Pyramid Lake (Nevada) was near historical lows and was dry for three days prior to sampling; and d) Eagle Lake had a water level of 5,091.5 ft (1,551.9 m), which was within 0.5 ft (0.2 m) of the lowest level recorded in 1935.

The categorization of a river mouth is a combination of the geographical location (e.g., the Columbia River) and evidence of the presence of the water mass the river flows into (e.g., salt water intrusion from the near-shore bay). The inflow of smaller rivers, streams, and creeks into a larger water body (e.g., Ward Creek flowing into Lake Tahoe) are not classified as river mouths, but rather as parent water mass modification from a creek inflow. The creek designation is to ensure the understanding that the inflow volume is small, but the expectation is changes in water properties are nonetheless discernible, because of the enhanced sensitivity (e.g., VSR) of the methods used.

The categorization of a water mass subjected to pollution, an atypical bloom, HAB, or an invasive species relied principally on eutrophic chlorophyll concentrations plus historical reporting or on-site local representatives. The latter are frequently present at boat ramps to oversee measures to mitigate health concerns or prevent the spread of invasive species, which is presently a significant and escalating problem throughout the western U.S.

2.6 NOMAD Archive

The NASA bio-Optical Marine Algorithm Dataset (NOMAD) v2.a (Werdell and Bailey 2005) is a small, quality controlled subset of a larger data repository established early in the Sea-viewing Wide Field-of-view Sensor (SeaWiFS) satellite mission (Hooker and Esaias 1993) called the SeaWiFS Bio-optical Archive and Storage System (SeaBASS) and is described by Hooker et al. (1994). The NOMAD database does not include applicable $a_{\mathrm{CDOM}}(440)$ measurements with contemporaneous UV and NIR spectral end members as used in this study. Consequently, the Hooker et al. (2013) algorithmic approach, which is based on UV and NIR end members, cannot be evaluated.

The NOMAD database, however, does include $K_d(\lambda)$ with legacy visible (VIS) wavelengths plus matching dissolved (Gelbstoff) spectral absorption coefficient at 443 nm, $a_g(443)$, which is functionally equivalent to $a_{\mathrm{CDOM}}(443)$ following Röttgers and Doerffer (2007). The consequences of the 3 nm shift in $a_g(443)$ with respect to $a_{\mathrm{CDOM}}(440)$ are considered negligible for a generalized inquiry involving legacy optical data, because the fixed wavelengths involved have 10 nm bandwidths and there are multiple sources of uncertainties in the derived data products of equal or greater importance (Hooker et al. 2013), e.g., pressure tares, aperture offsets, dark currents, wave focusing, etc.

2.7 Alternative Classification Scheme

The subcategory scheme (Sect. 2.5) included factual knowledge combined with careful inspection regarding one or more significant constituents or stressors influencing water mass complexity. Previous studies applied fuzzy c-means (FCM) classification to ocean color algorithm development (Moore et al. 2001) or uncertainty estimation (Moore et al. 2015), thereby demonstrating the usefulness of the FCM over a crisp or hard (e.g., k-means) classification. Application of the FCM classification is to log-transformed $K_d(\lambda)$ data based on the Calinski and Harabasz (1974) index (Matsuoka et al. 2013). All of the reported FCM classes are based on the water masses sampled herein, except the first class includes the influence of the K_d values of pure seawater, denoted K_w (Morel and Maritorena 2001).

3. Results

All data not categorized as one of the 15 subcategories prior to data collection (Sect. 2.5) are retained in the open ocean, coastal zone, and inland waters primary categories to yield the following number of validation quality observations, respectively: 190, 223, and 196. This new filtered data set of 609 observations is reasonably balanced, because each primary category contains approximately 200 observations. These data are used to initially evaluate the global applicability of the original Hooker et al. (2013) $a_{\mathrm{CDOM}}(440)$ algorithm. The comparison of the new validation quality data (i.e., data not part of the 15 subcategories) with respect to the original algorithm is presented in Fig. 3. The enhanced global sampling of the new data set with respect to Hooker et al. (2013) yields the following distinctions: a) the addition of lacustrine water bodies (including Lake Shikotsu in Hokkaido, Japan), which almost span the entire three decades of dynamic range (e.g., Crater Lake to Pinto Lake); and b) the expansion of the dynamic range to include clearer and more turbid water masses, which also means deeper and shallower waters.

Given the diversity of sampling in this study, it is unlikely that the more than three decades in $a_{\mathrm{CDOM}}(440)$ data (Fig. 3) might nonetheless exhibit similar chemical composition, i.e., the quantity of CDOM varies but the type remains constant. To reject the latter, the Fig. 3 CDOM spectral slope, S, values were compared to a global compilation of marine CDOM estimates from over 500 oceanographic campaigns (Aurin et al. 2018). The new study reported a comparable spectral range (350–600 nm) and a median S of 0.0167 nm^{-1} with a range of 0.0090–0.0208 nm^{-1}. In comparison, Fig. 3 data have a similar median of 0.0175 nm^{-1} and a slightly larger range of 0.0095–0.0410 nm^{-1}. In addition, Grunert et al. (2018) compiled similar statistics by marine province, with a comparable result in spectral range (350–550 nm) and an approximate range in S of 0.01–0.05 nm^{-1}. Consequently, the range of S presented in this study is comparable to similar recent global analyses spanning the majority of marine waters.

The new validation quality data significantly adhere to the original algorithm, as evidenced by how the red, green, and blue circles in Fig. 3 are well contained within the approximately ±15% gray boundaries that denote the dispersion in the original algorithm. The category that spans the largest percentage of the dynamic range for both axes is the inland waters data, although the coastal zone is somewhat similar because of the clear Hawaiian coastal waters that were sampled. The open ocean category has the smallest dynamic range, but this does not diminish the importance of this category because it represents the greatest surface area and volume of water on the planet.

The new data values, V_n, are compared to the corresponding algorithm value, with the latter being the reference values, V_r, in the comparison calculation. The relative percent difference (RPD) between the new data and the algorithm is computed as $\mathrm{RPD} = 100\,(V_n - V_r)/Vr$, and is expressed as a percent. The average RPD for all the new data is 0.02%, i.e., the new data show a negligible bias with respect to the original algorithm. The absolute percent difference (APD), which provides an estimate of the dispersion of differences between the new data and the algorithm, is the absolute value of the RPD. The average APD value for all the new data is 3.9%, i.e., the new validation quality data are usually to within 5% of the original algorithm (as visually confirmed by Fig. 3).

3.1 Drought-Stricken, Alkaline, Hypersaline, and Refilled Lakes

The new lacustrine data are presented in Fig. 4. Data from the hypersaline and alkaline (endorheic) lakes do not conform with the algorithm. Drought-stricken lakes exhibit a wider range of departure, with the most significant occurring for the most depleted water bodies, e.g., Lake Almanor and Shasta Lake. Endorheic drought-stricken lakes, e.g., Eagle Lake and Pyramid Lake, are the most extreme. Refilled lakes also do not conform with the algorithm, and refilled drought-stricken lakes exhibit an increase in CDOM and turbidity, e.g., Shasta Lake and Pyramid Lake.

The refilled lakes in Fig. 4 are frequently more different with respect to the algorithm than hypersaline or alkaline lakes, especially in terms of turbidity as determined by the K_d ratio. This is because some of the refilled lakes are overfilled, wherein the shore of the lake extends beyond the normal acreage of the lake (e.g., Washoe Lake and Little Washoe Lake). In overfilled lakes, land that is not normally flooded is added as new lake bottom, and the new acreage is a source of atypical constituents, either in composition or concentration (e.g., atypical constituents from land-use activities can be added when a lake overfills, because these activities are not possible in the water mass). The refilling of a normally dry endorheic basin, e.g., White Lake, wherein the flood waters and the reclaimed lake bottom provide the maximum areal and volumetric source of dissolved and suspended constituents results in some of the most extreme results, both in terms of turbidity and with respect to the algorithm.

The discharge from overfilled reservoirs also has significant deviations with respect to the algorithm. Thermalito Afterbay receives the discharge from Lake Oroville and was sampled after the overfilling of the parent water mass during the drought-breaking California 2016–2017 winter. Data were obtained in two locations, with higher CDOM data obtained in a shallow marsh. The ability to distinguish small localized differences establishes the sensitivity of the methods used herein. For example, the three refilled Shasta Lake samplings in Fig. 4 were conducted in different locations subjected to the inflow of a creek, as well as a large floating and partially submerged debris field.

3.2 River Mouth, Resuspension, and Ice Edge Proximity

The inflow of dissolved and suspended constituents to a parent water mass is explored further by considering a variety of sources that can add to water mass complexity. The new data are shown in Fig. 5 and were obtained in river mouths, water bodies with known suspension or visible resuspension, plus samples obtained closest to or farthest from the ice edge within an oceanic ice field. Water bodies with known suspension or visible resuspension are primarily from tidal and riverine flows, which are shown in Fig. 5 as triangles. Almost all of the resuspension data were obtained at peak tidal flow to ensure safe navigation in the necessarily shallow waters. The Akkeshi Bay data were obtained the day after the passage of typhoon Vongfong, wherein the shallow bay waters were a distinctly different color than normal. The Sacramento River data were obtained after heavy rains, wherein the boat ramp to be used was closed due to flood waters. The difference between the flooded Sacramento River with respect to the inland riverine data in Fig. 3 not in flood conditions (i.e., as conservative water masses), shows the classification of the Sacramento River is appropriate and the subjective classification approach has merit.

The resuspension data in Fig. 5 also include Bear Lake (which straddles the Utah-Idaho border) plus the effects of a large ship docking in the shallow RWC Channel with the aid of a tugboat. The latter involved the churning up of bottom material that significantly changed the color of the water. The resuspension sampling occurred shortly after the ship was docked. The Bear Lake scattering anomaly is created primarily through ground water seepage, which is rich in calcium carbonate particles (Davis and Milligan 2011). The ground water is nutrient poor and the small amount of riverine input to the lake is through a swamp and wetlands, wherein plants consume the nutrients and sediments settle out. Consequently, the Bear Lake data represent a significant clear-water scattering anomaly.

The other clear-water data in Fig. 5 were obtained principally in proximity to Arctic oceanic ice fields, and are distinguished as being closer to, or farther from, the ice. These data are displaced above or below the algorithm, respectively, even in the turbid waters of Kotzebue Sound. The majority of these data were obtained using a small boat launched from a larger ice breaker, so the data obtained closer to the ice are as safely close to the ice as possible while beyond the shading of the water mass by the ice field. The classification of *closer to* and *farther from* is qualitative and in complicated ice fields misclassification is possible. The Fig. 5 data show only two *farther from* points that are likely not classified correctly, and this category is the most vulnerable to a qualitative error.

In regards to the resuspension data, which all cluster below the algorithm in Fig. 5, river mouth data are the opposite—the data cluster above, but the number of such observations is much smaller. The reduced number is due to the difficulty of operating a trailered small boat in a shallow river, and then safely navigating the vessel out into the river mouth through a frequently narrow channel, wherein the higher sea state of the coastal ocean can be significantly amplified and boat traffic can make station work hazardous. Within the plume of a river mouth, two usually rather different water masses meet and mix over short time scales. Under those conditions, short-term deviations, with respect to the algorithm, can emerge and that is what is shown in the Fig. 5 data.

3.3 Atypical Blooms, Invasive Species, and Harbors

The presence of an atypical bloom, particularly a HAB, was anticipated to create additional optical complexity, because one or more significant stressors are frequently involved, e.g., an overabundance of nutrients, which can be anthropogenic in origin (Heisler et al. 2008). In a generic context, an atypical bloom includes the concentration of biomass to artificial levels (Kudela et al. 2015), perhaps due to local weather, (e.g., advective processes from winds and waves). Invasive species and harbors were also expected to increase optical complexity.

The new data obtained in harbors and water bodies experiencing an invasive species, or atypical bloom, including a HAB, are shown in Fig. 6. Some of these data could have had two classifications. For example, the Tahoe Keys and Tahoe Yacht Club were both infested with an invasive aquatic plant. Limited presence in one and mechanical removal in the other implied a harbor subcategory was appropriate. The Willamette River data were from an invasive aquatic species area (Bierly et al. 2015) and the algorithmic relationship is opposite the Lake Tahoe harbors. The latter suggests the Lake Tahoe harbor classifications, which cluster with the other harbors are likely appropriate.

Almost all harbors exhibit elevated $a_{\mathrm{CDOM}}(440)$ values with respect to the adjacent parent water mass, e.g., Chula Vista, Treasure Isle, San Leandro, America's Cup, etc. The relationship of harbors with respect to the algorithm has few extreme values, which is expected because harbors exchange water with the parent water mass. San Leandro and El Granada have the largest expression, but San Leandro is in a heavily urbanized area immediately south of Oakland International Airport in the San Francisco Bay area, so significant anthropogenic sources are anticipated.

Like some coastal harbors, El Granada vessels are moored in an inner shallow harbor protected by an outer deeper area, both with perimeter breakwaters and narrow channels. The two harbor areas cannot exchange water completely, i.e., a portion of the water volume is trapped during each tidal cycle, and are more turbid than the parent water mass, Half Moon Bay. The inner harbor is a likely and persistent anthropogenic source with a longer residence time, so it is anticipated to have an $a_{\mathrm{CDOM}}(440)$ value exceeding the neighboring bay. The increased residence time and reduced exchange rates through the narrow channels are a possible mechanism to increase $a_{\mathrm{CDOM}}(440)$. Other harbors wherein a protected moorage has elevated $a_{\mathrm{CDOM}}(440)$ include Las Vegas (Lake Mead) and Crescent City.

The HAB data in Fig. 6 were frequently obtained opportunistically and, thus, were not necessarily from the peak of the phenomenon. Also, a bloom is heterogeneous and navigation within the bloom is mostly based on visual observations, so the relationship with respect to the algorithm is not always extreme. The Monterey Bay HAB data are the most extensive, because there was the opportunity for scheduling some of the data collection during a time period when a HAB was likely to occur. In all cases, a HAB observation has a larger K_d ratio than the algorithm predicts, and this is principally caused by an increase in the $K_d(320)$ value, i.e., increased attenuation in the UV, which might indicate the presence of mycosporine-like amino acids (Jessup et al. 2009 and Kwon et al. 2018).

An atypical bloom is primarily a combination of local reporting, and a heterogeneous eutrophic water mass, i.e., chlorophyll concentration exceeds $1\,\mathrm{mg\,m^{-3}}$, with some water bodies having concentrations greater than $10\,\mathrm{mg\,m^{-3}}$. Consequently, the lack of sophistication and specificity related to explaining this subcategory does not exclude a simpler explanation. For example, local wind conditions could elevate the values associated with a typical bloom into atypical concentrations. This phenomenon was observed in more than one lake, e.g., Pyramid Lake and Upper Klamath Lake. The majority of the atypical blooms are in rather close agreement with the algorithm.

3.4 Wetlands, Pollution, and Water Mass Modifiers

The new data obtained in wetlands or polluted waters are presented in Fig. 7. The former are almost all marsh grass except two, which are labeled as to their types. The two unlabeled at the top of the plot are from Cutoff Slough in California and are marsh grass. All wetlands exhibit the same relationship, that is, they are all displaced above the algorithm, although four are in rather close agreement with the algorithm. The polluted water masses are associated with agricultural (Upper Klamath Lake and Upper Elkhorn Slough) or mining (Clear Lake) runoff, with the latter being the most severe. For both Upper Klamath Lake and Clear Lake, blue-green algae were plainly visible with extreme maximum chlorophyll concentrations of $1.117\,\mathrm{g\,m^{-3}}$ and $1.420\,\mathrm{g\,m^{-3}}$. The chlorophyll concentrations in Upper Elkhorn Slough were less, but are still extreme with a maximum value exceeding $100\,\mathrm{mg\,m^{-3}}$.

Figure 7 also includes examples of a small inflow from a creek or another source modifying the neighboring parent water mass. These data provide a measure of the sensitivity of the data acquisition, processing, and analysis techniques used herein. Although other sensitivity examples are documented above, e.g., the distinction between sampling closer to, or farther from, the ice edge (Fig. 5), the Fig. 7 examples span diverse spatial scales, e.g., creek inflows, a fish kill in the Salton Sea, and a large floating and partially submerged debris field in Shasta Lake. In all cases, the anticipated algorithmic relationships appear different than the parent water mass. The water properties of the creek inflow are not known, because access to the source from a small boat was problematic.

The generalized properties of the inflowing creek waters, determined visually, are as follows: a) the Lake Tahoe inflow was turbid, milky meltwater from snow and ice melting on shorelands; b) the Shasta Lake inflow was from rocky, tree-covered terrain and was significantly clearer than the lake water (the water pooled into a small pond before flowing into the lake and was easily observed); c) the Donner Lake inflow was from a rocky, tree-lined canyon; d) the Mono Lake inflow was across a mostly barren, rock-strewn shore with loose soil and was notably brown compared to the green lake; and e) the Pinto Lake inflow was from a densely vegetated buffer zone adjacent to farmland. The displacement of the modified waters with respect to the parent water mass are in keeping with these observations, i.e., the waters subjected to turbid or clear inflows had larger or smaller $K_d(320)/K_d(780)$ ratios, respectively.

3.5 Alternative Spectral End Members

The end-member wavelengths used in alternative $K_d(\lambda_1)/K_d(\lambda_2)$ ratios, hereafter $\Lambda_{\lambda_2}^{\lambda_1}$, follow the combinations first used by Hooker et al. (2013), i.e., the UV-NIR Λ_{710}^{340} pair, as well as the VIS Λ_{670}^{412} pair. Shortly after the start of this study, C-OPS system 021 was upgraded (Table 1), so the Λ_{875}^{313} pair is also available and provides the widest spectral span (562 nm) between end members. A plot of the end-member combinations is presented in Fig. 8, which also includes the linear fits and the root mean square error (RMSE) of the data with respect to the fits. The data in Fig. 8 are only those observations provided in Fig. 3, i.e., all 15 subcategories established in Sect. 2.5 (Figs. 4–7) are excluded. The consequences of using an increasing number of all the observations are presented in Sect. 4.

The fits in Fig. 8 show the end-member pair with the best accuracy is Λ_{780}^{320}, although the Λ_{875}^{313} and Λ_{710}^{340} fits are to within the calibration uncertainty of the radiometers plus inevitable environmental variance, i.e., to within 5%. The slope of the Λ_{780}^{320} fit is to within 1.1% of the original Hooker et al. (2013) algorithm ($y = 0.2556x - 0.0030$). As end-member wavelengths are brought spectrally closer together, the variance increases and reaches a maximum for the Λ_{670}^{412} pair, which degrades accuracy (RMSE generally increases with decreasing spectral separation of the end members). The Λ_{875}^{313} RMSE is a little larger than for Λ_{780}^{320} and a little less than for Λ_{710}^{340}. The fewer number of Λ_{875}^{313} data creates gaps in the data distribution, which partially explains why these data do not yield the lowest RMSE.

The Λ_{875}^{313} Fig. 8 data show the variance also increases after the transition from more turbid to clear waters, i.e., $a_{\mathrm{CDOM}}(440) = 0.02\,\mathrm{m}^{-1}$, and continues to increase with increasing water clarity. The larger variance as a function of water clarity is caused by the increasing importance of wave-focusing effects coupled with increasing NIR attenuation. Both problems are tractable for Λ_{780}^{320}, but contribute to the difficulty of deriving data products and ultimately producing a stable Λ_{875}^{313}. The increased Λ_{710}^{340} and Λ_{670}^{412} variances are not restricted to the problems described for Λ_{875}^{313}. As end members are brought spectrally closer together, the range of expression available to distinguish two similar but optically different water masses decreases. Consequently, choosing the extrapolation interval is more sensitive to small changes in the parameters that ultimately determine the fit for the extrapolation interval. For legacy end members, clear waters have a lesser range of expression and turbid waters have the greatest, so this problem decreases as turbidity increases, which is seen in the Λ_{710}^{340} and Λ_{670}^{412} Fig. 8 data.

3.6 Legacy Data Archive

From the full set of 4,459 NOMAD stations, 227 include Λ_{670}^{412} end members and $a_g(443)$ observations, hereafter $a_{\mathrm{CDOM}}(440)$, but 2 are duplicates. Application of Λ_{670}^{412} data to the corresponding algorithm in Fig. 8 results in 13 observations with negative (predicted) $a_{\mathrm{CDOM}}(440)$ values, which are removed to leave 212 unique stations. This process demonstrates how end-member algorithms can be used to quality assure optical data in archives. Of the 212 retained NOMAD stations, 189 are located within the Chesapeake Bay and its outflow into the southern Mid-Atlantic Bight, i.e., 89% of the data are from a restricted geographic area. For the remainder, 13 are off the mouth of Delaware Bay, 9 are from Massachusetts Bay, and 1 is in the open ocean northeast of South America. The southern Mid-Atlantic Bight and parts of Massachusetts Bay were part of the data used by Hooker et al. (2013), so data from these areas are anticipated to be compliant with the end-member algorithm. The average depth of Chesapeake Bay is relatively shallow (6.4 m) with a signifiant portion (over 24%) less than 2 m deep. Given the extensive contribution of rivers, tributaries, and tides to bay dynamics, resuspension of material is anticipated to be a source of bias in optical properties with respect to end-member algorithms for some bay stations (as shown in Fig. 5).

The retained NOMAD data are separated into two regimes: north Chesapeake Bay (NCB) and all other water masses, which consist of 106 stations for each. The dividing line for the NCB is the latitude of the Wicomico River in the Maryland Eastern Shore (slightly north of the Potomac River mouth). The separation is arbitrary and is used to compare the 106 NCB observations from NOMAD with 174 C-OPS K_d ratios and $a_{\mathrm{CDOM}}(440)$ data pairs obtained in the NCB (not shown in Fig. 2), albeit at different times and locations than the NOMAD data. During data collection, the C-OPS sampling was with system 021 (Table 1) and included notations about *in situ* conditions useful for establishing a resuspension subcategory, but the procedures predated and were not as rigorous as Sect. 2.5.

The C-OPS and NOMAD data plotted in Fig. 9 show general agreement (linearity) of the NOMAD data with respect to the algorithm, which independently confirms the Hooker et al. (2013) algorithmic approach (and as evaluated in more detail herein). Within the narrower turbidity range of the NOMAD and C-OPS NCB data without likely resuspension, there is improved agreement. The C-OPS NCB resuspension data appear properly categorized, because of their relationship to the algorithm (Fig. 5). There is evidence the C-OPS data considered free of resuspension effects nonetheless include some resuspension (e.g., some solid circles in Fig. 9 extend into the open circles as part of shallow-to-deep transects, thereby indicating the transect point in which resuspension effects were assumed absent was likely premature). The NOMAD data exhibit a higher variance with respect to the algorithm, which results in an increased RMSE of 37.8% (or 44.1% if the 13 omitted observations are included) compared to the 6.2% value determined with C-OPS data (Fig. 8). The more extreme NOMAD values suggest a subcategorization methodology that could be applied to archival data would improve agreement with the algorithm (already demonstrated with the removal of 13 observations using the Λ^{412}_{670} algorithm in Fig. 8).

If the NOMAD data are partitioned into turbid and clear subsets, using $a_{\mathrm{CDOM}}(440) > 0.2$ and $a_{\mathrm{CDOM}}(440) \leq 0.2$ as thresholds, respectively, the fit equation for the turbid Λ^{412}_{670} data is $y = 0.3437x - 0.2404$. The slope of this turbid NOMAD fit is similar to the corresponding end-member fit presented in Fig. 8 for which $y = 0.3504x - 0.1033$, and agrees to within 1.9%. The fit for the clear NOMAD data, however, is $y = 0.0758x + 0.0648$, which is significantly different at the 78.4% level. The 13 NOMAD stations that were not retained out of the original 225 NOMAD stations were in clear waters, which is another indicator that the NOMAD data in clearer waters are problematic.

With respect to the algorithm, the increased bias, variance, and 13 negative derived values obtained with NOMAD data in clearer waters suggests the legacy data are degraded by sampling artifacts. Example degradations in legacy data include wave-focusing effects (because of the slower sampling rates), coarser VSR (because of faster descent rates), and deeper extrapolation intervals (because of near-surface data loss from large vertical tilts and large aperture depth offsets). Although some legacy data problems are absent from C-OPS data (e.g., because there is no righting moment when C-OPS sampling begins and C-PrOPS stabilizes the planar orientation of all apertures, some aspects of these limitations are present in the Fig. 8 data, but they are not significant, i.e., they result in a small increase in variance, which slightly degrades algorithm performance. Inclusion of the C-OPS NCB data without resuspension to deriving the Λ^{412}_{670} algorithm (Fig. 8) results in rather small changes to the fit coefficients. The slope is to within 4.3% and the intercept is to within 4.6% (both within the net 5% uncertainty for calibration and environmental variance).

3.7 Objective versus Subjective Classification

The data set established herein has an extensive number of observations directly suitable for validation exercises (Figs. 3 and 9) plus 15 subcategories (Sect. 2.5) of more complex and, thus, potentially (but not automatically) problematic, water bodies (Figs. 4–7), with the latter determined subjectively. The combination yields 16 categorizations of data spanning an arguably *global* sampling of open ocean, coastal zone, and inland water masses in terms of a generalized perspective of the dynamic range in water properties (Figs. 3–8). The NOMAD search (Sect. 3.6), however, showed archival data provided a significantly less global data set in terms of spectral expanse and dynamic range as used herein. Archival data usually do not include a subcategory parameter for the observations, e.g., NOMAD has no applicable keyword. Although some subcategories could be determined from geolocation, temporal, and survey information (e.g., a harbor, wetland, alkaline lake, etc.), other influences are not usually established without an observer (e.g., atypical bloom or resuspension caused by a vessel). Consequently, a subcategorization based on the optical measurements alone might be advantageous to the validation process, particularly for archival data.

The subcategory approach is evaluated using $K_d(\lambda)$ spectra for the aforementioned 16 subcategories of data, which can be described objectively based on spectral shapes and magnitudes. A small number of observations are excluded to ensure consistency in the determination of all $K_d(\lambda)$ values, e.g., the White Lake data had estimated values in the UV domain, Bear Lake is a unique scattering anomaly created by calcium carbonate particles, ship-induced resuspension is anthropogenic in origin, etc. With the additional restriction of wavelength commonality spanning 320–780 nm (Table 1), a total of 1,171 spectra are used for the objective classification analysis.

Application of the FCM classification to log-transformed $K_d(\lambda)$ data for the 16 subcategorizations successfully classifies all $K_d(\lambda)$ spectra into five classes based on the Calinski and Harabasz (1974) index (Sect. 2.7). Spectral shapes, as well as their magnitudes, uniquely vary between the five classes and span a continuum of water masses from oceanic and lacustrine case-1 to extreme case-2 inland waters. The continuum of water mass composition is summarized by the centered spectra for the five FCM classes shown in Fig. 10, where N_i is the class number set by index i. The diversity achieved in sampling lacustrine water masses is revealed in Fig. 10 by the range of $K_d(\lambda)$ spectra for the example drought-stricken and refilled lakes shown, which are compared with Crater Lake and K_w. The refilled lakes are shown to emphasize the complex relationships presented in Figs. 4–7 are not detectable by $K_d(\lambda)$ alone, but require an understanding of the applicable end-member ratio and the $a_{\mathrm{CDOM}}(440)$ value.

The two lakes at the bottom and top of the dynamic range in Fig. 10, are Crater Lake and White Lake, respectively, with the latter having the largest displacement with respect to the algorithm in Figs. 4–7. Using the inverse of $K_d(\lambda)$ as a proxy for the vertical scale that must be properly sampled (i.e., a sufficient number of observations must be obtained within the vertical scale to derive data products), the vertical scale in Fig. 10 ranges from meters (Crater Lake) to millimeters (White Lake), with the latter only being possible with unprecedented VSR. The shorter wavelengths for White Lake are shown as "estimated" because these wavelengths had to be processed with individual extrapolation intervals to provide the K_d estimates, which is not in keeping with all the other data products.

The average $K_d(313)$ value obtained for deep water (586 m) sampling in Crater Lake was 0.072 m^{-1} and the coefficient of variation from the six casts was 1.1%. According to Morel et al. (2007) Crater Lake is "extremely" clear water, and the CV shows exceptional reproducibility, i.e., 1% radiometry. Note that as shown in Fig. 3, Crater Lake was sampled twice. A station with higher $a_{\mathrm{CDOM}}(440)$ values was conducted in shallower water above submerged moss that grows in large, dense mats. The constituent properties of the water were anticipated to be influenced by the moss matts, and the $a_{\mathrm{CDOM}}(440)$ values are elevated with respect to the deeper station.

The proportionate composition of each FCM class in Fig. 10 as a function of the original subjective subcategories, wherein contributions less than 5% are reported but not considered significant, is presented in Table 2. Using Fig.

10 and Table 2, the corresponding principal class characteristics are as follows:

N_1 The smallest contributor to the dynamic ranges, although arguably accounting for most of the pixels in a global CDOM image, with the $K_d(\lambda)$ maximum in the NIR and the minimum in the blue domain (400–490 nm). The spectral shape is consistent with typical case-1 waters (Morel and Maritorena 2001). The proportional makeup is dominated by the validation quality subcategory (Figs. 3 and 8) at 83% with farther from ice data (Fig. 5), which are almost exclusively from slightly modified case-1 waters, contributing an additional 14%.

N_2 The $K_d(\lambda)$ maximum in the NIR is similar to case-1 waters, and the minimum is shifted to longer wavelengths (490–565 nm) from case-1 modifications, principally from proximity to ice effects. The validation quality proportion decreases to 66% and is supplanted with proximity to ice subcategories (Fig. 5).

N_3 The $K_d(\lambda)$ minimum is near the middle of the green domain (555–565 nm) due to increasing optical complexity as case-2 constituents appear in larger proportions. The UV domain values are the same as, or slightly lower than, the NIR domain. The validation quality proportion decreases to 39% while case-2 subcategories increase significantly, i.e., resuspension, drought-stricken and refilled lakes, harbors, and HABs.

N_4 The $a_{\mathrm{CDOM}}(440)$ dynamic range is established with the $K_d(\lambda)$ minimum shifted into the green-red domains (555–625 nm) and maximum values in the UV exceeding the NIR. Optical complexity reaches a maximum, because all 16 categorizations contribute at the 1% level or more. The validation quality proportion is the most abundant, but is decreased to 36%. The resuspension, harbors, and refilled lakes subcategories provide net increases in case-2 waters with additional extreme contributions from alkaline and hypersaline lakes.

N_5 Extreme waters that only extend the optical dynamic range, with the $K_d(\lambda)$ minimum in the NIR domain (710 nm), and maximum values compared to N_1–N_4 that peak in the UV. The resuspension subcategory is dominant at 38%, followed by drought-stricken lakes at 14%, and the validation quality subcategory is reduced to 13%. The remaining principal contributors are wetlands, refilled lakes, and polluted water bodies.

The decrease in the percent composition of the validation quality data as a function of increasing class number (N_1–N_5) is an indicator of the difficulty of validating an algorithm within increasingly complex waters. The recurring contribution of a relatively small number of principal subjective subcategories to the gradient in optical complexity confirms the subcategory approach has merit and reveals the cause-and-effect relationships of the subcategories.

Table 2 also indicates the resolution or granularity for the 15 subcategories of more complex water masses was more nuanced than required. For example, alkaline and hypersaline lakes could be one subcategory, as could refilled and drought-stricken lakes. Combining subcategories does not ensure an eventual convergence with the objective FCM classification, because the latter partitions the processes present in the subcategories into varying degrees of contribution for each identified class. This partitioning involves both direct and indirect evidence, which is perhaps best realized with the original granularity of the 15 subcategories as revealed by considering resuspension processes.

The direct evidence of resuspension is provided by the resuspension subcategory (created for this phenomenon). Indirect resuspension is present or likely in multiple subcategories, however. For example, melt water releases particles at the ice edge, bottom deposits in harbors are stirred up by boat propellers, wind and rain redeposit exposed bottom material into drought-stricken lakes, refilled lakes contain suspended material from riverine inflow, tidal and wave action suspend material in shallow wetlands, etc. Table 2 data show the percent contribution of direct and indirect suspension increases with increasing class number N_1 (17%) to N_3 (46%) to N_5 (77%), so a significant part of the cascade towards complexity is correlated with water masses not evolving conservatively, e.g., due to resuspension.

A small number of principal subcategories (N_p) significantly determine the $K_d(\lambda)$ classification spectra, although the paucity of observations for some subcategories, in part opportunistic or planned depending on the subcategory, is a largely unknown mitigating factor. Nonetheless, if refilled and drought-stricken lakes plus hypersaline and alkaline lakes, are respectively considered one subcategory rather than two, which appears reasonable based on Fig. 4, five or less principal subcategories determine 97%, 92%, 85%, 80%, and 96% of the composition for N_1–N_5, respectively.

Without the subcategory scheme, the importance of direct or indirect resuspension would not have emerged with the clarity provided in Table 2. The composition of the FCM classes confirm the applicability of subcategories while revealing which ones are important to each class and which ones are marginally or not important. For the latter, the parent water mass modifier is not significant, as expected, because this small subcategory was created to demonstrate the sensitivity of the *in situ* optical and laboratory methods. Other subcategories might appear insignificant because of the difficulty of obtaining the sample (e.g., river mouth data are underrepresented) or the paucity of opportunities to collect a sample (e.g., local restrictions to prevent the spread of invasive species).

Because metadata to subcategorize NOMAD data are unavailable, the objective classification scheme was applied to NOMAD and to determine the number of data in each class, as follows: N_1 6, N_2 13, N_3 135, N_4 49, N_5 0, plus 9 were unclassified. All the turbid data are in classes N_3 and N_4; the clear data are in classes N_1 and N_2, plus 9 are unclassified. This means the slope of the aforementioned clear partition (Sect. 3.6) had 19 points that were classified and 9 that were not, and likely accounts for the poor performance with respect to the algorithm. It also suggests the measurements were the issue with the NOMAD data, because some spectra could not be classified.

4. Discussion

The optical data herein had a near-surface VSR less than 1 mm, which allowed data products spanning 313–875 nm while encompassing a global perspective of water masses as described by approximately three decades of generalized water properties. The validation approach was based on the concept that water masses evolving conservatively (i.e., free from stressors that might alter the natural range in the gradient of a constituent) are suitable for validating the original Hooker et al. (2013) inversion algorithm for deriving $a_{\mathrm{CDOM}}(440)$ from $K_d(\lambda)$ spectral end members.

The identification of 15 subcategories that were likely not evolving conservatively yielded 609 validation quality data points spanning extremely clear to highly turbid waters sampled within the open ocean, coastal zone, and inland waters. The new data adhered to the original Λ^{320}_{780} algorithm (Fig. 3) with an RPD of 0.02% and an APD of 3.86%. Alternative spectral end members (e.g., Λ^{313}_{875}, Λ^{340}_{710}, and Λ^{412}_{670}) had increasingly larger RMSE values, but were to within the calibration uncertainty of the radiometers plus inevitable environmental variance (i.e., a net uncertainty to within 5%) except for the narrowest spectral span of legacy end members, which was 6.2% for Λ^{412}_{670} (Fig. 8).

Although no data archive exists with the spectral and spatial coverage used herein, NOMAD Λ^{412}_{670} data showed general agreement (linearity) with respect to the original Hooker et al. (2013) algorithm, and independently confirmed the algorithmic approach. There was also general agreement between the NOMAD NCB data and the corresponding C-OPS NCB data without likely resuspension, wherein the C-OPS NCB resuspension data appeared properly categorized, because of their relationship with the algorithm (Fig. 5).

The near-surface VSR is quantitative evidence of the successful mitigation of a variety of sampling difficulties (e.g., large aperture offsets, righting moment instabilities, wave focusing effects, etc.). The high VSR achieved with C-OPS resulted in an increased sensitivity for deriving data products in turbid water masses with vertical scales on the order of millimeters (Fig. 10). This sensitivity allowed small localized changes in a parent water mass to be distinguished (Figs. 5 and 7). The laboratory analyses were similarly sensitive, so the optical determinations of in-water constituents using field measurements were commensurately as effective as laboratory analyses. The ability to distinguish small differences in water properties ensured the discrimination of 15 subcategories, wherein each represented a more complex water mass not evolving conservatively and not automatically used for validation.

Plots of the $a_{\mathrm{CDOM}}(440)$ and Λ^{320}_{780} relationships for the 15 subcategories of more complex water bodies revealed some data were significantly different with respect to the original algorithm and others that were not (Figs. 4–7). The wide range in complexity is largely the result of the substantial effort that was made to adopt a global perspective and sample the greatest diversity of water bodies possible, some of which were very difficult to access and sometimes required shoreline launches of the optical instrumentation that C-PrOPS made possible, e.g., severely drought-stricken and refilled lakes (high and dry, flooded, or debris-blocked boat ramps) plus hypersaline and soda lakes (Mono Lake, Salton Sea, Borax Lake, and Soda Lake had no boat ramps).

The accuracy of the algorithm as a function of including increasing proportions of the 15 subcategories of more complex water masses not necessarily appropriate for validation exercises (Figs. 4–7), because they were likely not evolving conservatively, is explored by expanding the 609 validation quality observations of Λ_{780}^{320} end members (Figs. 3 and 8) to include the following subcategories (hereafter referred to as the second data set): inflows to a parent water mass that are not hypersaline or drought-stricken lakes, closer to or farther from ice edge proximity, river mouth, resuspension (but not including Bear Lake and the ship-induced RWC Channel resuspension), atypical blooms, HABs, and wetlands. This second data set has 930 observations, the linear fit is $y = 0.2511x - 0.0046$, the RMSE is 5.7%, and the new slope is to within 1.7% of the original value presented by Hooker et al. (2013), i.e., $y = 0.2556x - 0.0030$.

The reason the slope of the second data set is not significantly different than the original fit coefficients is the data that were added are situated above and below the distribution of the validation quality data set, as shown in Figs. 4 and 5. The C-OPS NCB data without resuspension cluster on or below the algorithmic relationship (Fig. 9). If these data are added to the second data set used to derive the Λ_{780}^{320} algorithm (the 320 nm and 780 nm wavelengths were always part of system 021 as shown in Table 1), the resulting slope and intercept is $y = 0.2561x - 0.0076$, which is to within 2.0% of the slope determined for the second data set ($y = 0.2511x - 0.0046$).

If a third data set is created by adding drought-stricken and refilled lakes to the second data set, but not including White Lake (which had some estimated data products), this third data set has 1,044 observations. The linear fit is $y = 0.2249x + 0.0044$, the RMSE is 6.8%, and the new slope is reduced (as expected, because the added data are all below the algorithmic relationship and primarily turbid (Fig. 4), but still within 10.4% of the original value presented by Hooker et al. (2013). The exclusion of White Lake, as well as hypersaline, alkaline, and polluted water bodies, from the third data set is for a practical reason: they are, or have significant characteristics of, extreme water masses and White Lake rarely exists.

There are other lakes presented herein that can be considered extreme, e.g., other endorheic lakes (e.g., Pyramid Lake, Eagle Lake, etc.), plus shallow lakes in high wind areas wherein bottom material is resuspended on a near-continuous basis (e.g., Washoe Lake and Little Washoe Lake). If all subcategories, except extreme lacustrine water bodies, are used to create a fourth data set it has 1,086 observations, i.e., almost 90% of the 1,230 maximum and 93% of the data used in Table 2 to create the five objective FCM classifications. The linear fit of this fourth more comprehensive data set is $y = 0.2379x - 0.0049$, the RMSE is 6.2%, and the new slope is not reduced as much as in the third data set (as expected) and is to within 6.9% of the original value presented by Hooker et al. (2013).

A fifth data set, using classes N_1–N_4 spans the entire $a_{\mathrm{CDOM}}(440)$ dynamic range (Table 2). This alternative for an arguably *global* algorithm yields a linear fit of $y = 0.2317x - 0.0053$, the RMSE is 5.3%, and the slope is to within 9.3% of the original Hooker et al. (2013) value. This result is significantly similar to the subjective results of the fourth data set (the slopes agree to within 2.7%). Consequently, the robustness of the algorithm is directly supported by the combination of subjective and objective classifications, with the latter using FCM.

With the exception of the third data set that added primarily turbid data exclusively below the algorithmic relationship (drought-stricken and refilled lakes), all of the results from the expanded data sets are rather indistinguishable from the original Λ_{780}^{320} fit provided by Hooker et al. (2013), wherein $y = 0.2556x - 0.0030$, or the separate validation quality data set presented in Fig. 8, for which $y = 0.2583x - 0.0053$. Ignoring the third data set, the x-intercept for the expanded data sets is approximately equal to what can be expected for pure water, i.e., 0.02. The close agreement of the various expanded data sets with adherence to the pure-water limit is another significant confirmation of the algorithmic approach using spectral end members.

The robustness is further established by creating a so-called *universal algorithm*, which is assumed to mean that any water mass wherein an optical profiler can be deployed is expected to be part of the evaluation of the end-member approach. In this case, the universal algorithm is constructed from all the data from all subcategories, and is distinguished from the global perspective in that the universal dataset includes a far greater proportion of complex water types than exist globally. The linear fit of this universal data set is $y = 0.2206x + 0.0088$, the RMSE is 7.5%, and the new slope is to within 13.7% of the original value presented by Hooker et al. (2013). The removal of data that are, or have significant characteristics of, extreme water masses improves performance. For example, if hypersaline, alkaline, and polluted lakes are removed, the linear fit of this sixth comprehensive data set is $y = 0.2250x + 0.0024$, the RMSE is 6.8%, and the new slope is to within 12.0% of the original value presented by Hooker et al. (2013).

This study used three different laboratory methods to determine $a_{\mathrm{CDOM}}(440)$ from water samples, and seven different optical instrument suites to determine $K_d(\lambda)$. Despite the agreement between the Λ_{780}^{320} fits and their x-intercepts, the possibility the results are the result of an unidentified stochastic process has not been addressed. The latter is not likely for the optical data, because the radiometers were calibrated at one facility and deployed with strict adherence to the Protocols using the same acquisition software. Furthermore, data products were derived using the same processing software with one operator, and the variance in optical data products is shown in Figs. 3–8.

The capabilities of the above- and in-water radiometers for C-OPS systems 021 (with C-PrOPS) and 039 (without C-PrOPS) were intercompared to next-generation *hybridspectral* instruments (Hooker et al. 2018c) containing a hyperspectral detector system plus 18 fixed wavelengths (system 038 in Table 1). The comparisons showed an agreement of 4.2–4.8%, which is to within the calibration uncertainty (2.3–2.7%) plus natural variability (i.e., a net uncertainty to within 5%), as long as the stability threshold for a backplane without thrusters or the noise limit of the hyperspectral sensor was not exceeded (Hooker et al. 2018b and 2018c, respectively).

Although field data demonstrate small changes in parent water mass modifications can be discriminated (Figs. 5 and 7), the laboratory methods and instruments were not systematically intercalibrated to establish an overall uncertainty. In addition, the laboratory methods included different temperature controls and storage procedures for the water samples, as well as application of null-point corrections in different spectral ranges. A lack of systematic intercalibration is not a significant detraction, because it is a common difficulty when combining observations from databases (e.g., NOMAD data), and methods exist to nonetheless determine the efficacy of the combined data.

Following the technique established by Matsuoka et al. (2017), a statistical sensitivity analysis is used to examine the uncertainty of the combined $a_{\mathrm{CDOM}}(440)$ values from the three different laboratory methods used in Figs. 3–8. Briefly, for each optical and water sampling, normality of distribution for $a_{\mathrm{CDOM}}(440)$ was created using the measured value as the mean (μ) and 7% of the measured value as the standard deviation (σ). Similarly, normality of distribution for Λ^{320}_{780} was created using the measured value as μ and 5% as σ. A lower σ percentage was used for the latter, because the original data set retained each optical cast whereas only one water sample was obtained, so the variability of the optical data was already significantly represented (Figs. 3–8).

For each original data pair, 10^5 variations were prepared for both Λ^{320}_{780} and $a_{\mathrm{CDOM}}(440)$. Of these data, 10^3 were randomly selected and the mean computed for each Λ^{320}_{780} and $a_{\mathrm{CDOM}}(440)$ pair. This exercise was repeated 10^3 times (bootstrap) and an overall μ, and σ were obtained. For the different combinations of applying $\mu \pm \sigma$ to bound the dispersion of the original data, the resulting linear fits showed the new algorithm slopes changed by 0.3–1.1% with respect to the Λ^{320}_{780} validation quality data set in Fig. 8. For all combinations of $\mu \pm \sigma$, the maximum RMSE was 1.1%, which is similar to the 1.2% value for Λ^{320}_{780} in Fig. 8. Consequently, the use of three different laboratory methods to determine $a_{\mathrm{CDOM}}(440)$, which were not intercalibrated, does not significantly influence the results presented here.

The state-of-the-art (approximately 1%) accuracy achieved with the Λ^{320}_{780} end members (Fig. 8) is, therefore, due to the strict adherence to sampling and processing protocols coupled with an unprecedented VSR for the optical data. This combination also resulted in all end-member pairs—including the legacy (VIS) Λ^{412}_{670} pair—having a superior accuracy compared to many common global inversion algorithms. For example, $a_{\mathrm{CDOM}}(\lambda)$ algorithms based on the water-leaving radiance, $L_W(\lambda)$, and its normalized forms have an RMSE exceeding 10% or more (Mannino et al. 2008 and 2014), and the ocean color (OC) chlorophyll a variants that use band ratios (O'Reilly et al. 1998 and 2000), typically exceed 20% or more while generally excluding complex and inland waters as used herein.

The robustness of the end-member approach, is further confirmed by how the expanded data sets (up to six), which had respectively increased amounts of water masses that were not evolving conservatively, nonetheless yielded fits with RMSE values outperforming the aforementioned $L_W(\lambda)$ algorithms. In regards to which of the variants best represents a global perspective, the $\Lambda_{\lambda_2}^{\lambda_1}$ pairs in Fig. 8 are considered appropriate. Screening of newly collected or archival data, e.g., NOMAD data (Sect. 3.7), with respect to a selected algorithm can be accomplished by initially flagging data points more than 12% from the expected relationship, and then more carefully examining those points using both objective and subjective criteria (based on available metadata) to determine whether the results are expected, or are more likely to indicate a problem with data collection procedures.

If an algorithm is to be applied to a water mass that is not evolving conservatively, an individualized relationship between $a_{\mathrm{CDOM}}(\lambda)$ and $\Lambda_{\lambda_2}^{\lambda_1}$ should be established, especially for an extreme water mass, e.g., drought-stricken. This maintains the accuracy of the global relationship while providing a mechanism for improving the study of nonconservative waters. Identification of water masses not evolving conservatively can likely be determined using $K_d(\lambda)$ values determined from in-water optical (e.g, C-OPS) data without a need for laboratory analyses (Fig. 9). With additional research to produce $K_d(\lambda)$ for all wavelengths, e.g., expanding upon Cao et al. (2014), the identification of water masses evolving conservatively or not can be made from above-water observations. This determination provides a sensitive indicator of water masses subjected to stresses influencing water quality, e.g., drought. The onset of next-generation satellites, e.g., the Japanese Second-generation Global Imager (SGLI) mission (Honda et al. 2012), offers unique opportunities for such monitoring because of the expanded spatial and spectral domain, with the latter allowing improved $\Lambda_{\lambda_2}^{\lambda_1}$ algorithm accuracy with respect to legacy missions, e.g., Moderate Resolution Imaging Spectroradiometer (MODIS). While planned high spectral resolution sensors, such as the Plankton, Aerosol, Cloud, ocean Ecosystem (PACE) and Surface Biology and Geology (SBG) missions, may support more sophisticated retrievals of parameters like CDOM, the simplified approach provided by end-member analysis can be used with both legacy and next-generation sensors, thereby providing continuity in space and time, as well as a capability to generate high quality in-water data with a simplified measurement approach (assuming strict adherence to the Protocols).

Acknowledgments

This work was principally supported by the National Aeronautics and Space Administration (NASA) as part of planning for the Aerosol, Cloud, Ecosystems (ACE) satellite remote sensing mission. The next-generation perspective benefitted from the anticipated calibration and validation activities of the SGLI, ACE, and PACE missions. Additional field sampling opportunities were provided by the following NASA projects: Hyperspectral Infrared Imager (HyspIRI); Coastal and Ocean Airborne Science Testbed (COAST); Coastal High Acquisition Rate Radiometers for Innovative Environmental Research (C-HARRIER); and Hybridspectral Alternative for Remote Profiling of Optical Observations for NASA Satellites (HARPOONS). Part of this study was supported by the Japan Aerospace Exploration Agency (JAXA) Global Change Observation Mission-Climate (GCOM-C) project to A. Matsuoka (principal investigator ER2GCF310). The high level of success achieved in the field work established a foundation of understanding that was the direct consequence of contributions from individuals who contributed unselfishly to the work involved (e.g., calibration, acquisition, processing, and sampling). The scientists included (alphabetically) J. Brown, B. Hargreaves, T. Hirawake, T. Isada, R. Lind, J. Morrow, K. Negrey, and J. Nishioka; their dedicated contributions are gratefully acknowledged.

References

Antoine, D., Hooker, S.B., Bélanger, S., Matsuoka, A., and Babin, M.: Apparent optical properties of the Canadian Beaufort Sea, part I: Observational overview and water column relationships. *Biogeosci.*, **10,** 4,493–4,509, 2013.

Aurin, D., Mannino, A., and Lary, D.J.: Remote Sensing of CDOM, CDOM Spectral Slope, and Dissolved Organic Carbon in the Global Ocean. Appl. Sci., **8,** 2687; doi:10.3390/app8122687, 2018.

Bierly, K., Burgess, P., Fetcho, K., Gregory, S., Graham-Hudson, B., Hudson, W., Hulse, D., Moore, K., Mullan, A., Wallick, R., Welch, D., and Wiley, P.: Tracking Progress in Restoring the Willamette River Floodplain. *Habitat Technical Team Pub.*, Oregon Department of Fish and Wildlife, Salem, Oregon, 13 pp, 2015.

Bodaker, I., Sharon, I., Suzuki, M.T., Feingersch, R., Shmoish, M., Andreishcheva, E., Sogin, M.L., Rosenberg, M., Maguire, M.E., Belkin, S. and Oren, A.: Comparative community genomics in the Dead Sea: an increasingly extreme environment. *The ISME journal*, **4,** 399–407, 2010.

Bushaw, K.L., Zepp, R.G., Tarr, M.A., Schultz-Jander, D., Bourbonniere, R.A., Hodson, R.E.m Miller, W.L., Bronk, D.A., and Moran, M.A.: Photochemical release of biologically available nitrogen from aquatic dissolved organic matter. *Nature*, **381,** 404–407, 1996.

Calinski, T., and Harabasz, J.: A dendrite method for cluster analysis. *Commun. Stat.*, **3,** 1–27, 1974.

Cao, F., Fichot, C.G., Hooker, S.B., and Miller, W.L.: Improved algorithms for accurate retrieval of UV/visible diffuse attenuation coefficients in optically complex, inshore waters. *Remote Sens. Environ.*, **144,** 11–27, 2014.

Davis, J., and Milligan, M.: Why is Bear Lake so Blue? And Other Commonly Asked Questions. *Public Information Series 96*, Utah Geological Survey, Salt Lake City, Utah, 42 pp, 2011.

Fellman, J.B., D'Amore, D.V., and Hood, E.: An evaluation of freezing as a preservation technique for analyzing dissolved organic C, N and P in surface water samples. *Sci. Total Environ.*, **392,** 305–312, 2008.

Gordon, H.R., and Ding, K.: Self shading of in-water optical instruments. *Limnol. Oceanogr.,* **37,** 491–500, 1992.

Grant, W.D.: Extremophilies—Alkaline environments and biodiversity. In: *Encyclopedia of Life Support Systems (EOLSS)*, C. Gerday, and N. Glansdorff, Eds., Eolss Publishers, Oxford, United Kingdom, 20 pp, 2006.

Green, S.A., and Blough, N.V.: Optical absorption and fluorescence properties of chromophoric dissolved organic matter in natural waters. *Limnol. Oceanogr.*, **39,** 1903–1916, 1994.

Grunert, B.K., Mouw, C.B., and Ciochetto, A.B.: Characterizing CDOM spectral variability across diverse regions and spectral ranges. *Global Biogeochem. Cycles*, **32,** 57–77, 2018.

Guarch-Ribot, A., and Butturini, A.: Hydrological conditions regulate dissolved organic matter quality in an intermittent headwater stream. From drought to storm analysis. *Sci. Total Environ.*, **571,** 1358–1369, 2016.

Heisler, J., Glibert, P., Burkholder, J., Anderson, D., Cochlan, W., Dennison, W., Gobler, C., Dortch, Q., Heil, C., Humphries, E., Lewitus, A., Magnien, R., Marshall, H., Sellner, K., Stockwell, D., Stoecker, D., and Suddleson, M.: Eutrophication and harmful algal blooms: A scientific consensus. *Harmful Algae*, **8,** 3–13, 2008.

Hencke, K., Hovland, E.K., Volent, Z., Pettersen R., Johnsen G., Moline M., and Sakshaug E.: Optical properties of CDOM across the Polar Front in the Barents Sea: Origin, distribution and significance. *J. Mar. Sys.*, **130,** 219–227, 2014.

Honda, Y., Kajiwara, K., Sharma, R., Ono, A., Imaoka, K., Murakami, H., Hori, M., Ono, Y., and Rostand, D.: Overview of GCOM-C1/SGLI science. *Proc. SPIE*, **8528,** doi: 10.1117/12.977772, 2012.

Hooker, S.B., and Esaias, W.E.: An overview of the SeaWiFS project. *Eos, Trans. Amer. Geophys. Union*, **74,** 241–246, 1993.

Hooker, S.B., McClain, C.R., Firestone, J.K., Westphal, T.L., Yeh, E-n., and Ge, Y.: The SeaWiFS Bio-Optical Archive and Storage System (SeaBASS), Part 1. *NASA Tech. Memo. 104566, Vol. 20,* S.B. Hooker and E.R.

Firestone, Eds., NASA Goddard Space Flight Center, Greenbelt, Maryland, 40 pp, 1994.

Hooker, S.B., Morrow, J.H., and Matsuoka, A.: Apparent optical properties of the Canadian Beaufort Sea, part II: The 1% and 1 cm perspective in deriving and validating AOP data products. *Biogeosci.,* **10,** 4,511–4,527, 2013.

Hooker, S.B., Zibordi, G., Berthon, J-F., D'Alimonte, D., Maritorena, S., McLean, S., and Sildam, J.: Results of the Second SeaWiFS Data Analysis Round Robin, March 2000 (DARR-00). *NASA Tech. Memo. 2001–206892, Vol. 15,* S.B. Hooker and E.R. Firestone, Eds., NASA Goddard Space Flight Center, Greenbelt, Maryland, 71 pp, 2001.

Hooker, S.B.: Mobilization Protocols for Hybrid Sensors for Environmental AOP Sampling (HySEAS) Observations. *NASA Tech. Pub. 2014–217518*, NASA Goddard Space Flight Center, Greenbelt, Maryland, 105 pp, 2014.

Hooker, S.B., Lind, R.N., Morrow, J.H., Brown, J.W., Suzuki, K., Houskeeper, H.F., Hirawake, T., and Maúre, E.R.: Advances in Above- and In-Water Radiometry, Vol. 1: Enhanced Legacy and State-of-the-Art Instrument Suites. *NASA Tech. Pub. 2018-219033/Vol. 1*, NASA Goddard Space Flight Center, Greenbelt, Maryland, 60 pp, 2018a.

Hooker, S.B., Lind, R.N., Morrow, J.H., Brown, J.W., Kudela, R.M., Houskeeper, H.F., and Suzuki, K.: Advances in Above- and In-Water Radiometry, Vol. 2: Autonomous Atmospheric and Oceanic Observing Systems. *NASA Tech. Pub. 2018-219033/Vol. 2*, NASA Goddard Space Flight Center, Greenbelt, Maryland, 69 pp, 2018b.

Hooker, S.B., Lind, R.N., Morrow, J.H., Brown, J.W., Kudela, R.M., Houskeeper, H.F., and Suzuki, K.: Advances in Above- and In-Water Radiometry, Vol. 3: Hybridspectral Next-Generation Optical Instruments. *NASA Tech. Pub. 2018-219033/Vol. 3*, NASA Goddard Space Flight Center, Greenbelt, Maryland, 39 pp, 2018c.

Houser, J.N.: Water color affects the stratification, surface temperature, heat content, and mean epilimnetic irradiance of small lakes. *Can. J. Fish. Aquat. Sci.*, **63,** 2,447–2,455, 2006.

Jessup, D.A., Miller, M.A., Ryan, J.P., Nevins, H.M., Kerkering, H.A., Mekebri, A., Crane, D.B., Johnson, T.A. and Kudela, R.M.: Mass stranding of marine birds caused by a surfactant-producing red tide. *PLoS One*, **4,** p.e4550, 2009.

Kudela, R.M., Palacios, S.L., Austerberry, D.C., Accorsi, E.K., Guild, L.S. and Torres-Perez, J.: Application of hyperspectral remote sensing to cyanobacterial blooms in inland waters. *Remote Sens. Environ.*, **167,** 196–205, 2015.

Kutser, T., Verpoorter, C., Paavel, B., and Tranvik, L.J.: Estimating lake carbon fractions from remote sensing data. *Remote Sens. Environ.*, **157,** 138–146, 2015.

Kwon, H.K., Kim, G., Lim, W.A. and Park, J.W.: In-situ production of humic-like fluorescent dissolved organic matter during *Cochlodinium polykrikoides* blooms. *Est. Coast. Shelf Sci.*, 203, 119–126, 2018.

Lee, Z., and Hu, C.: Global distribution of Case-1 waters: An analysis from SeaWiFS measurements. *Remote Sens. Environ.*, **101,** 270–276, 2006.

Mannino, A., Russ, M.E., and Hooker, S.B.: Algorithm development and validation for satellite-derived distributions of DOC and CDOM in the U.S. Middle Atlantic Bight. *J. Geophys. Res.,* **113,** C07051, doi:10.1029/2007JC004493, 2008.

Mannino, A., Novak, M., Hooker, S.B., Hyde, K., and Aurin, D.: Algorithm development and validation of CDOM properties for estuarine and continental shelf waters along the northeastern U.S. coast. *Remote Sens. Environ.*, **152,** 576–602, 2014.

Matsuoka, A., Bricaud, A., Benner, R., Para, J., Sempéré, R., Prieur, L., Bélanger, S., and Babin, M.: Tracing the transport of colored dissolved organic matter in water masses of the Southern Beaufort Sea: relationship with hydrographic characteristics. *Biogeosci.,* **9,** 925–940, 2012.

Matsuoka, A., Hooker, S.B., Bricaud, A., Gentili, B., and Babin, M.: Estimating absorption coefficients of colored dissolved organic matter (CDOM) using a semi-analytical algorithm for Southern Beaufort Sea waters: application to deriving concentrations of dissolved organic carbon from space. *Biogeosci.*, **10,** 917–927, 2013.

Matsuoka, A., Babin, M., Doxaran, D., Hooker, S.B., Mitchell, B.G., B/'elanger, S., and Bricaud, A.: A synthesis of light absorption properties of the Arctic Ocean: application to semianalytical estimates of dissolved organic carbon concentrations from space. *Biogeosci.*, **11,** 3,131–3,147, 2014.

Matsuoka, A., Boss, E., Babin, M., Boss, L-K., Hafez, M., Chekalyuk, A., Proctor, C.W., Werdell, J., and Bricaud, A.: Pan-Arctic optical characteristics of colored dissolved organic matter: Tracing dissolved organic carbon in changing Arctic waters using satellite ocean color data. *Remote Sens. Environ.*, **200,** 89–101, 2017.

Moore, T.S., Campbell, J.W., and Feng, H.: A fuzzy logic classification scheme for selecting and blending satellite ocean color algorithms. *IEEE Trans. Geosci. Remote Sens.*, **39,** 1,764–1,776, 2001.

Moore, T.S., Campbell, J.W., and Feng, H.: Characterizing the uncertainties in spectral remote sensing reflectance for SeaWiFS and MODIS-Aqua based on global *in situ* matchup data sets. *Remote Sens. Environ.*, **159,** 14–27, 2015.

Mopper, K., Zhou, X., Kieber, R.J., Kieber, D.J., Sikorski, R.J., and Jones, R.D.: Photochemical degradation of dissolved organic carbon and its impact on the oceanic carbon cycles. *Nature*, **353,** 60–62, 1991.

Morel, A.: "Optical Properties of Pure Water and Pure Sea Water." In: *Optical Aspects of Oceanography*, Jerlov, N.G., and E.S. Nielsen, Eds., Academic Press, New York, 1–24., 1974

Morel, A., and Maritorena, S.: Bio-optical properties of oceanic waters: a reappraisal. *J. Geophys. Res.,* **106,** 7,163–7,180, 2001.

Morel, A., Gentili, B., Claustre, H., Babin, M., Bricaud, A., Ras, J., and Tièche, F.: Optical properties of the "clearest" natural waters. *Limnol. Oceanogr.*, **52,** 217–229, 2007.

Morrow, J.H., Hooker, S.B., Booth, C.R., Bernhard, G., Lind, R.N., and Brown, J.W.: Advances in Measuring the Apparent Optical Properties (AOPs) of Optically Complex Waters. *NASA Tech. Memo. 2010–215856,* NASA Goddard Space Flight Center, Greenbelt, Maryland, 80 pp, 2010.

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.: Aquatic color radiometry remote sensing of coastal and inland waters: Challenges and recommendations for future satellite missions. *Remote Sens. Environ.*, **160,** 15–30, 2015.

Mueller, J.L.: "Overview of Measurement and Data Analysis Protocols" In: G.S. Fargion and J.L. Mueller, Ocean Optics Protocols for Satellite Ocean Color Sensor Validation, Revision 2. *NASA Tech. Memo. 2000–209966,* NASA Goddard Space Flight Center, Greenbelt, Maryland, 87–97, 2000.

Mueller, J.L.: "Overview of Measurement and Data Analysis Protocols." In: J.L. Mueller and G.S. Fargion, Ocean Optics Protocols for Satellite Ocean Color Sensor Validation, Revision 3, Volume 1. *NASA Tech. Memo. 2002–210004/Rev3–Vol1,* NASA Goddard Space Flight Center, Greenbelt, Maryland, 123–137, 2002.

Mueller, J.L.: "Overview of Measurement and Data Analysis Methods." In: Mueller, J.L., and 17 Coauthors, Ocean Optics Protocols for Satellite Ocean Color Sensor Validation, Revision 4, Volume III: Radiometric Measurements and Data Analysis Protocols. *NASA Tech. Memo. 2003–211621/Rev4–Vol.III*, NASA Goddard Space Flight Center, Greenbelt, Maryland, 1–20, 2003.

Mueller, J.L., and Austin, R.W.: Ocean Optics Protocols for SeaWiFS Validation. *NASA Tech. Memo. 104566, Vol. 5,* S.B. Hooker and E.R. Firestone, Eds., NASA Goddard Space Flight Center, Greenbelt, Maryland, 43 pp, 1992.

Mueller, J.L., and Austin, R.W.: Ocean Optics Protocols for SeaWiFS Validation, Revision 1. *NASA Tech. Memo. 104566, Vol. 25,* S.B. Hooker, E.R. Firestone, and J.G. Acker, Eds., NASA Goddard Space Flight Center, Greenbelt, Maryland, 66 pp, 1995.

Nelson, N.B., and Siegel, D.A.: The global distribution and dynamics of chromophoric dissolved organic matter. *Annu. Rev. Mar. Sci.*, **5,** 447–476, 2013.

Nelson, N.B., and Siegel, D.A., Carlson, C.A., Swan, C.M., Smethie Jr., W.M., and Khatiwala, S.: Hydrography of chromophoric dissolved organic matter in the North Atlantic. *Deep-Sea Res. I*, **54,** 710–731, 2007.

O'Reilly, J.E., Maritorena, S., Mitchell, B.G., Siegel, D.A., Carder, K.L., Garver, S.A., Kahru, M., and McClain, C.: Ocean Color chlorophyll algorithms for SeaWiFS. *J. Geophys. Res.,* **103,** 24,937–24,953, 1998.

O'Reilly, J.E., and 24 Coauthors: SeaWiFS Postlaunch Calibration and Validation Analyses, Part 3. *NASA Tech. Memo. 2000–206892, Vol. 11*, S.B. Hooker and E.R. Firestone, Eds., NASA Goddard Space Flight Center, Greenbelt, Maryland, 9–23, 2000.

Palmer, S.C., Kutser, T., and Hunter, P.D.: Remote sensing of inland waters: challenges, progress and future directions. *Remote Sens. Environ.*, **157,** 1–8, 2015.

Röttgers, R., and Doerffer, R.: Measurements of optical absorption by chromophoric dissolved organic matter using a point-source integrating-cavity absorption meter. *Limnol. Oceanogr. Methods*, **5,** 126–135, 2007.

Smith, R.C., and Baker, K.S.: The analysis of ocean optical data. *Ocean Optics VII,* M. Blizard, Ed., *SPIE,* **478,** 119–126, 1984.

Tanaka, K., Takesue, N., Nishioka, J., Kondo, Y., Ooki, A., Kuma, K., Hirawake, T., and Yamashita, Y.: The conservative behavior of dissolved organic carbon in surface waters of the southern Chukchi Sea, Arctic Ocean, during early summer. *Sci. Rep.*, **6,** 34123, 2016.

Twardowski, M.S., Boss, E., Sullivan, J.M., and Donaghay, P.L.: Modeling the spectral shape of absorption by chromophoric dissolved organic matter. *Mar. Chem.*, **89,** 69–88, 2004.

Vazquez, E., Amalfitano, S., Fazi, S., and Butturini, A.: Dissolved organic matter composition in a fragmented Mediterranean fluvial system under severe drought conditions. *Biogeochemistry,* **102**, 59–72, 2011.

Werdell, P.J., and Bailey, S.W.: An improved in-situ bio-optical data set for ocean color algorithm development and satellite data product validation. *Remote Sens. Environ.*, **98,** 122–140, 2005.

Yamashita, Y., and Tanoue, E.: Basin scale distribution of chromophoric dissolved organic matter in the Pacific Ocean. *Limnol. Oceanogr.*, **54,** 598–609, 2009.

Yamashita, Y., Nosaka, Y., Suzuki, K., Ogawa, H., Takahashi, K., and Saito, H.: Photobleaching as a factor controlling spectral characteristics of chromophoric dissolved organic matter in open ocean. *Biogeosci.*, **10,** 7,207–7,217, 2013.

Yapiyev, V., Sagintayev, Z., Inglezakis, V., Samarkhanov, K. and Verhoef, A.: Essentials of endorheic basins and lakes: A review in the context of current and future water resource management and mitigation activities in Central Asia. *Water,* **9,** 798, 2017.

Zaneveld, J.R.V., Boss, E., and Barnard, A.: Influence of surface waves on measured and modeled irradiance profiles. *Appl. Opt.*, **40,** 1,442–1,449, 2001.

Table 1. The nominal fixed wavelengths in nanometers, all with 10 nm bandwidths, for each optical profiling system as distinguished by serial number. The number of casts obtained is differentiated between backplanes without (C-OPS) and with (C-PrOPS) digital thrusters (note system 021 had both), wherein boldface numbers indicate the 15 wavelengths common to all profiling systems, e.g., 320 nm and 780 nm.

Optical Profiling System							*Number of Casts*	
010	014	021	021†	034	038§	039	C-OPS	C-PrOPS
	313		313			313	57	709
320	320	320	320	320	320	320	**497**	**733**
340	340	340	340	340	340	340	**497**	**733**
380	380	380	380	380	380	380	**497**	**733**
395	395	395	395	395	395	395	**497**	**733**
412	412	412	412	412	412	412	**497**	**733**
443	443	443	443	443	443	443	**497**	**733**
465	465	465	465	465		465	497	709
490	490	490	490	490	490	490	**497**	**733**
510	510	510	510	510	510	510	**497**	**733**
532	532	532	532	532	532	532	**497**	**733**
555	555	555	555	555	555	555	**497**	**733**
560		560					356	0
				565			84	0
	589		589		589	589	57	733
625	625	625	625	625	625	625	**497**	**733**
665		665		665			440	0
670	670	670	670	670	670	670	**497**	**733**
683	683	683	683	683	683	683	**497**	**733**
710	710	710	710	710	710	710	**497**	**733**
780	780	780	780	780	780	780	**497**	**733**
					820		0	24
	875	875	875	875	875	875	470	733

† Upgraded with C-PrOPS, a conductivity sensor, and new wavelengths. § Hypbridspectral and equipped with C-PrOPS.

Table 2. The objective FCM classification of the data in Figs. 3–8 with a few omissions to ensure consistent data quality (e.g., White Lake, Bear Lake, ship-induced resuspension, etc.). The five classes N_i, where i is the class number index, are shown with the number of $K_d(\lambda)$ spectra, N_s, within each class in parentheses, as well as the percent composition of the original 16 subjective subcategories equalling or exceeding a 1% contribution threshold for each class. The principal subcategories in each class, i.e., the most numerous (approximately 5% contribution or more), are shown in bold typeface. After the subcategories, the number of principal subcategories (N_p) and the number of all subcategories (N_a) with a 1% composition or more are summarized in slanted typeface followed by the percent composition from all the principal subcategories. The extent of the dynamic range in percent for the optical and biogeochemical data are shown in the last two lines as a function of applying successive class numbers.

Original Subcategory Name (N_s)	*Class Number and Composition* N_1 (244)	N_2 (263)	N_3 (305)	N_4 (265)	N_5 (94)
Validation Quality	**83%**	**66%**	**39%**	**36%**	**13%**
Farther from Ice	**14**	**10**	1	2	
Closer to Ice	2	**11**	1	1	
Resuspension	1	1	**10**	**16**	**38**
Refilled Lake		1	**9**	**7**	**10**
Drought-Stricken Lake		4	**8**	3	**14**
Harbor (or Marina)			**11**	**8**	4
Harmful Algal Bloom		1	**8**	2	1
Wetland (or Marsh)			3	3	**11**
Polluted Water Mass				4	**10**
Alkaline Lake			2	**5**	
Hypersaline Lake			2	**5**	
River Mouth			3	3	
Atypical Bloom		3	1	2	
Parent Water Mass Modifier		2	3	2	
Invasive Species				2	
N_p (N_a) Subcategories	*2(4)*	*3(9)*	*6(14)*	*6(16)*	*6(8)*
N_p Percent Composition	**97%**	**87%**	**85%**	**77%**	**96%**
$K_d(320)/K_d(780)$ *Extent*	2%	8%	23%	83%	100%
$a_{\mathrm{CDOM}}(440)$ *Extent*	3%	9%	36%	100%	100%

Figure Captions

Fig. 1. The next-generation C-OPS backplane with C-PrOPS (roll is the long axis and pitch is the short axis into or out of the page): a) irradiance cosine collector; b) radiometer bumper; c) array of 19 microradiometers; d) aggregator and support electronics; e) rotating V-block for pitch adjustment; f) two-point harness attachment; g) *hydrobaric* buoyancy chamber, which accommodates up to 3 compressible bladders; h) slotted flotation and i) bronze weights for buoyancy and roll adjustment; j) water temperature probe and k) pressure transducer port on the radiance end cap; l) conductivity sensor; m) electronics module; n) digital thruster (one of two, on each side); and o) thruster guard. The aggregator and support electronics control the 19 microradiometers as a single device. The side bumpers and thruster guards protect the radiometers and digital thrusters from unanticipated side impacts, respectively.

Fig. 2. The geographical distribution of the original Canadian Arctic and U.S. east coast data (open diamonds) used in Hooker et al. (2013) versus the new validation data (solid circles).

Fig. 3. The new validation quality data from the primary open ocean, coastal zone, and inland waters (blue, green, and red circles, respectively) categories, which are used to evaluate the original Hooker et al. (2013) algorithm (gray circles). The location names of a subset of observations are explicitly identified as a function of the approximately three decades of dynamic range in both axes. A ±7.5% dispersion is approximately represented by the larger algorithm symbol size, i.e., from one edge of a gray circle to the opposite edge represents a total of approximately 15.0% dispersion. The headwaters of the San Francisco Bay Redwood Creek (RWC) Channel, which is surrounded by wetlands, is the most turbulent coastal water mass and has the highest $a_{\mathrm{CDOM}}(440)$ value.

Fig. 4. The new data from lacustrine water bodies that were drought-stricken (magenta diamonds), refilled after drought (blue triangles), alkaline (red squares), or hypersaline (green circles) in relation to the original algorithm (gray circles). Each data point represents a single water sample with multiple optical casts for each, which results in a series of results (typically 3–6) along the x-axis.

Fig. 5. The new data obtained in river mouths (red circles), water bodies with known suspension or visible resuspension (green triangles), plus samples obtained closer to (magenta squares) or farther from (blue diamonds) the ice edge within an ice field. The San Francisco RWC Channel headland waters are depicted in Fig. 3 as are the Columbia and Umqua River data that are upstream of the river mouth.

Fig. 6. The new data obtained in a harbor (or marina), plus water bodies experiencing an invasive species, HAB, or atypical bloom.

Fig. 7. The new data obtained in a wetland or polluted water mass plus comparisons between a parent water mass and a creek inflow or another source of water properties.

Fig. 8. Four end-member algorithms to derive $a_{\mathrm{CDOM}}(440)$ from in-water optical observations with the accuracy of each estimated using RMSE statistics.

Fig. 9. The adherence of NOMAD archival data to the legacy (VIS) $K_d(412)/K_d(670)$ algorithm shown in Fig. 8 (gray solid circles) for the north Chesapeake Bay (NCB) and other Mid-Atlantic Bight locations (red and orange solid diamonds, respectively). The NOMAD NCB data are compared to C-OPS NCB data obtained at different times and locations with the latter separated into two categories, wherein one is likely subjected to bottom resuspension (light blue open circles) and the other is not (dark blue solid circles).

Fig. 10. The centered $K_d(\lambda)$ spectra of the five classes (N_i) determined from an objective FCM classification of the data presented in Figs. 3–8 with a few omissions for data consistency (e.g., White Lake, Bear Lake, ship-induced resuspension, etc.) and shown with respect to K_w. Example K_d spectra from drought-stricken and refilled lakes plus Crater Lake, obtained by averaging the results from multiple optical casts, are also shown to demonstrate the more than three decades of dynamic range in turbidity that was sampled. The shorter wavelengths for White Lake (open dark red squares) required individual wavelength processing to provide the estimated K_d values, whereas all other data were obtained with a single processing.

Figure 1

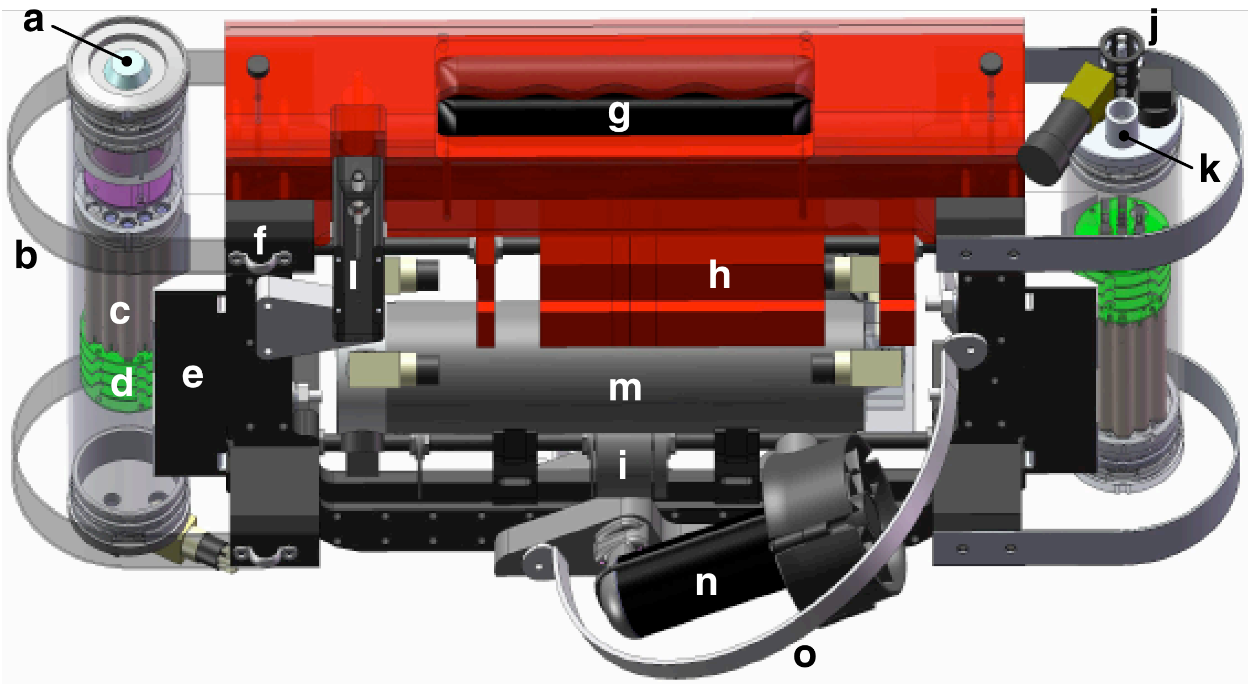


Figure 2

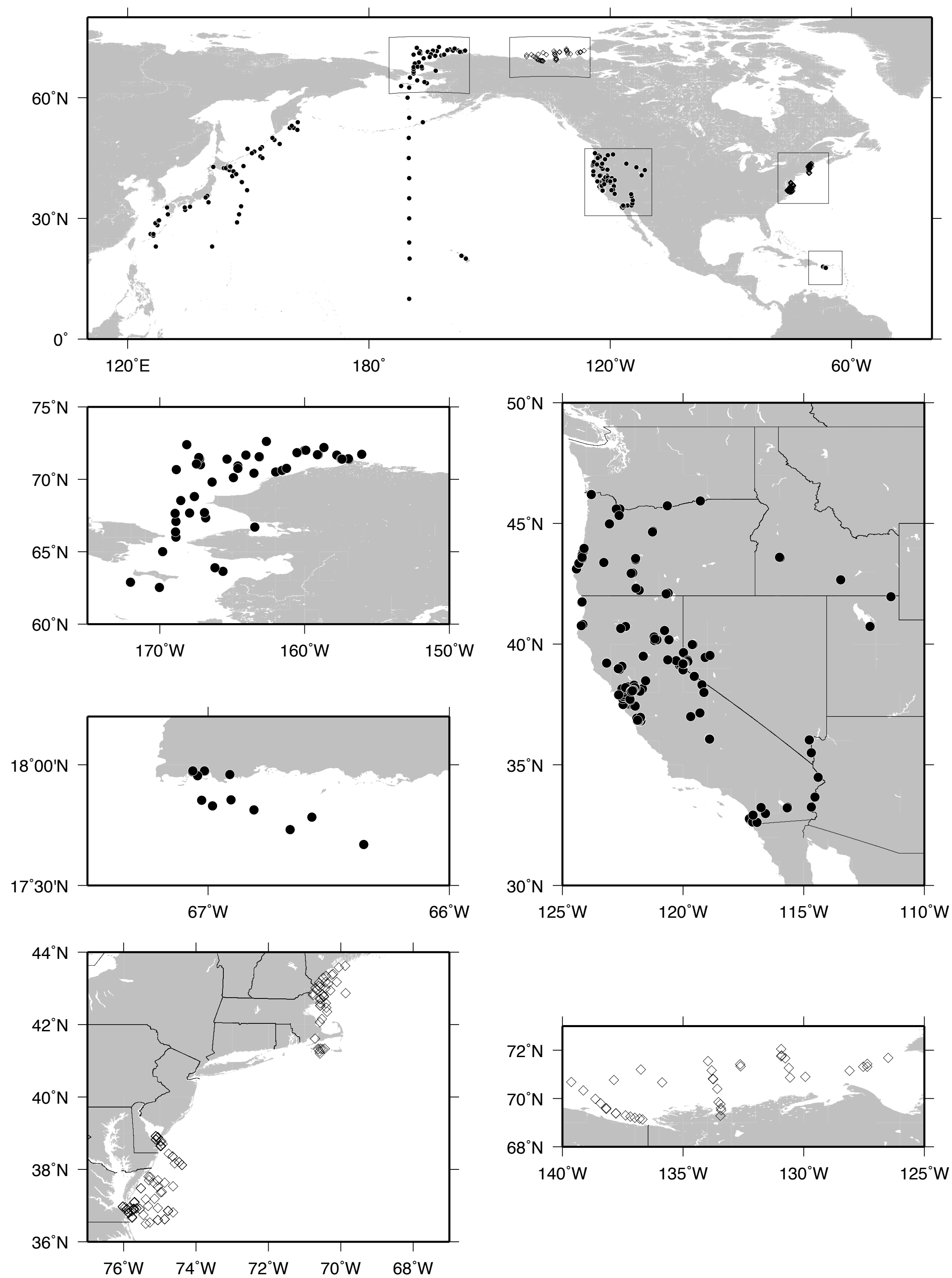

60˚N
30˚N
0˚
120˚E
180˚
120˚W
60˚W
75˚N
70˚N
65˚N
60˚N
170˚W
160˚W
150˚W
50˚N
45˚N
40˚N
35˚N
30˚N
125˚W
120˚W
115˚W
110˚W
18˚00'N
17˚30'N
67˚W
66˚W
44˚N
42˚N
40˚N
38˚N
36˚N
76˚W
74˚W
72˚W
70˚W
68˚W
72˚N
70˚N
68˚N
140˚W
135˚W
130˚W
125˚W

Figure 3

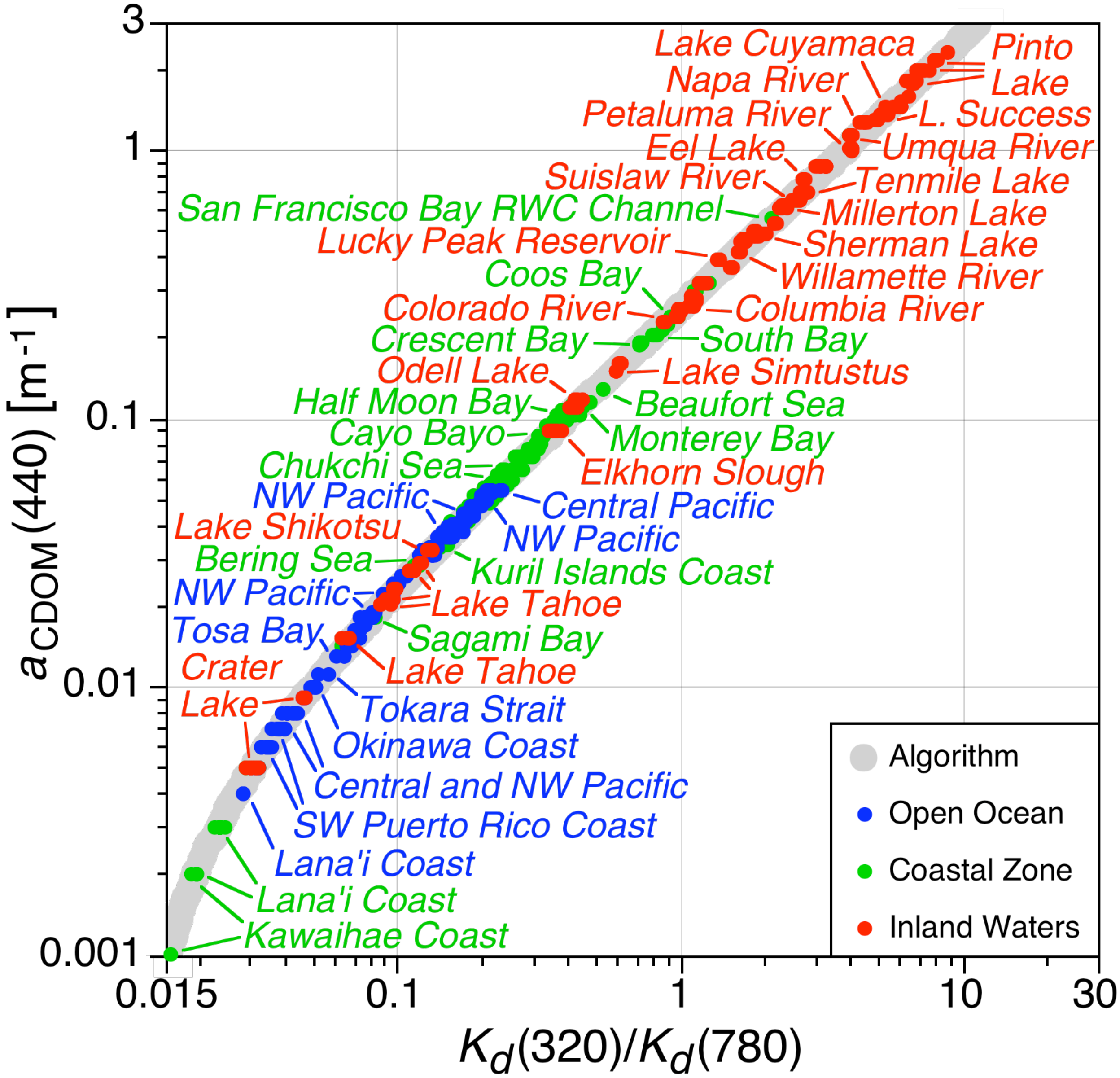
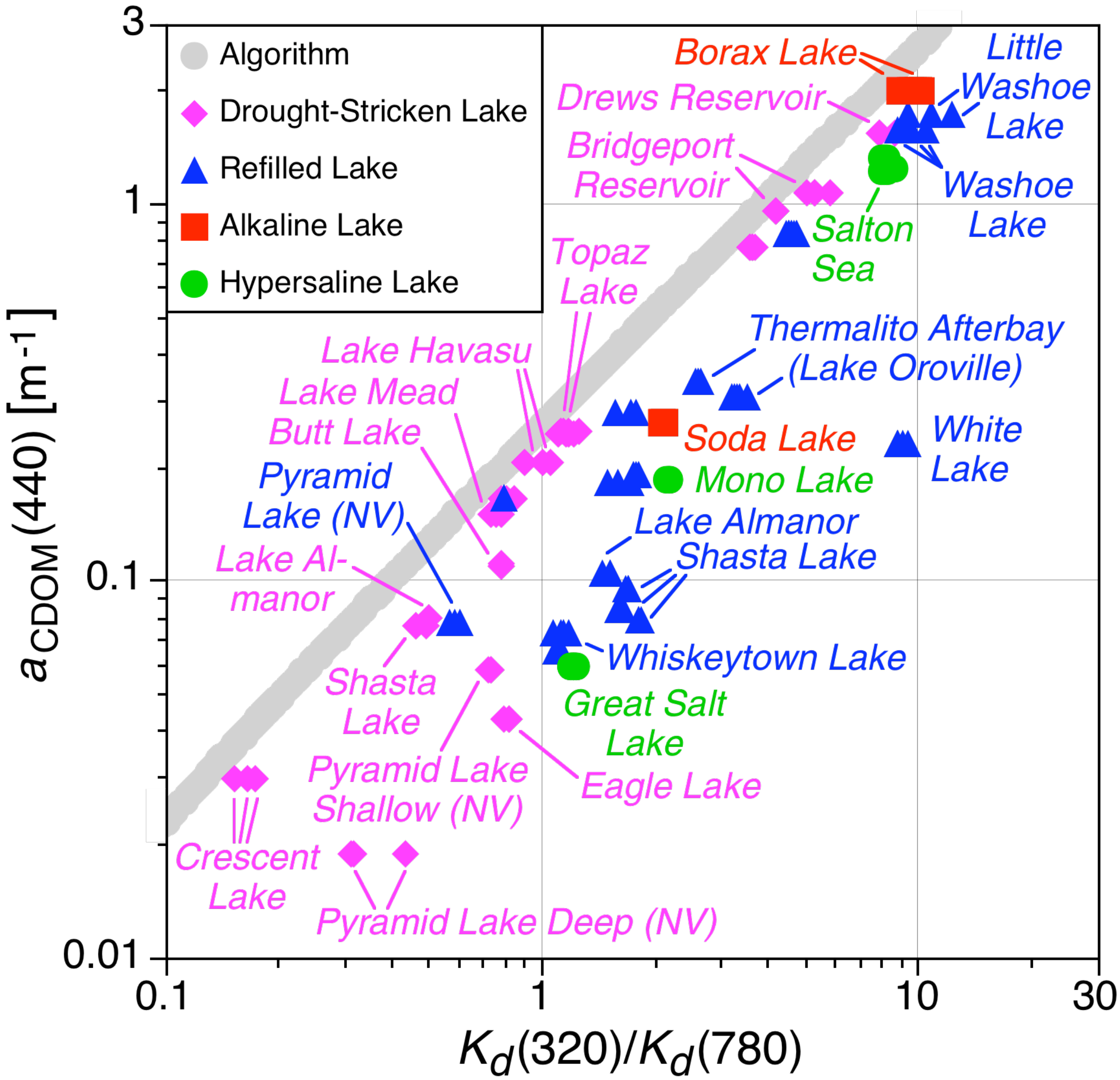

Lake Cuyamaca
Napa River
Petaluma River
Eel Lake
Suislaw River
San Francisco Bay RWC Channel
Lucky Peak Reservoir
Coos Bay
Colorado River
Crescent Bay
Odell Lake
Half Moon Bay
Cayo Bayo
Chukchi Sea
NW Pacific
Lake Shikotsu
Bering Sea
NW Pacific
Tosa Bay
Crater Lake
Pinto Lake
L. Success
Umqua River
Tenmile Lake
Millerton Lake
Sherman Lake
Willamette River
Columbia River
South Bay
Lake Simtustus
Beaufort Sea
Monterey Bay
Elkhorn Slough
Central Pacific
NW Pacific
Kuril Islands Coast
Lake Tahoe
Sagami Bay
Lake Tahoe
Tokara Strait
Okinawa Coast
Central and NW Pacific
SW Puerto Rico Coast
Lana'i Coast
Lana'i Coast
Kawaihae Coast
Algorithm
Open Ocean
Coastal Zone
Inland Waters
$a_{CDOM}(440)$ [m-1]
$K_d(320)/K_d(780)$
3
1
0.1
0.01
0.001
0.015
0.1
1
10
30

FIGURE 4



Algorithm
Drought-Stricken Lake
Refilled Lake
Alkaline Lake
Hypersaline Lake
Borax Lake
Little Washoe Lake
Drews Reservoir
Bridgeport Reservoir
Washoe Lake
Salton Sea
Topaz Lake
Thermalito Afterbay (Lake Oroville)
Lake Havasu
Lake Mead
Butt Lake
Soda Lake
White Lake
Mono Lake
Pyramid Lake (NV)
Lake Almanor
Shasta Lake
Lake Al-manor
Whiskeytown Lake
Shasta Lake
Great Salt Lake
Pyramid Lake Shallow (NV)
Eagle Lake
Crescent Lake
Pyramid Lake Deep (NV)
3
1
0.1
0.01
0.1
1
10
30
$a_{CDOM}(440)$ [m-1]
$K_d(320)/K_d(780)$

Figure 5

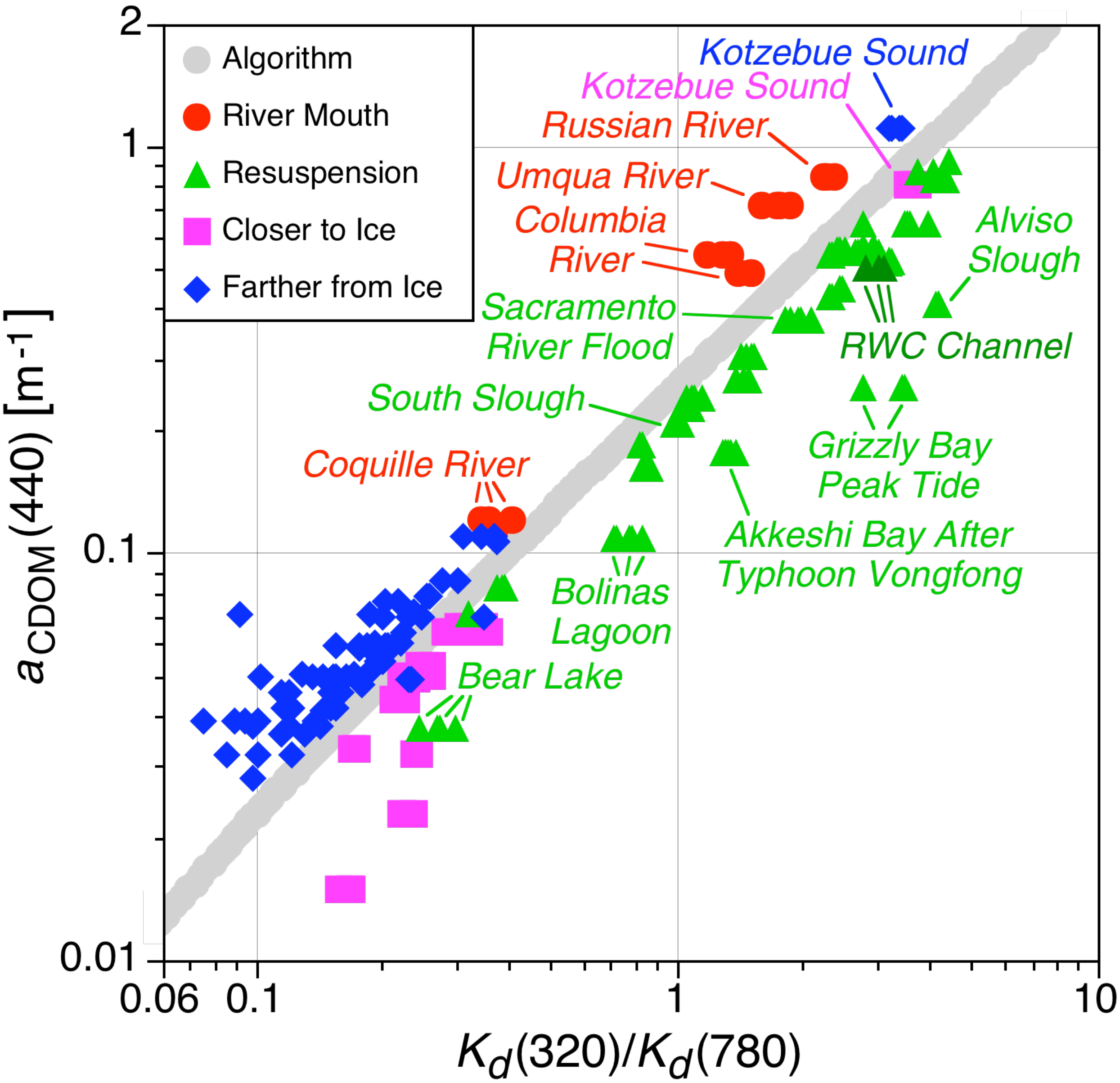

Algorithm
River Mouth
Resuspension
Closer to Ice
Farther from Ice
Kotzebue Sound
Kotzebue Sound
Russian River
Umqua River
Columbia River
Sacramento River Flood
South Slough
Coquille River
Alviso Slough
RWC Channel
Grizzly Bay Peak Tide
Akkeshi Bay After Typhoon Vongfong
Bolinas Lagoon
Bear Lake
2
1
0.1
0.01
0.06
0.1
1
10
$a_{CDOM}(440)$ [m^{-1}]
$K_d(320)/K_d(780)$

Figure 6

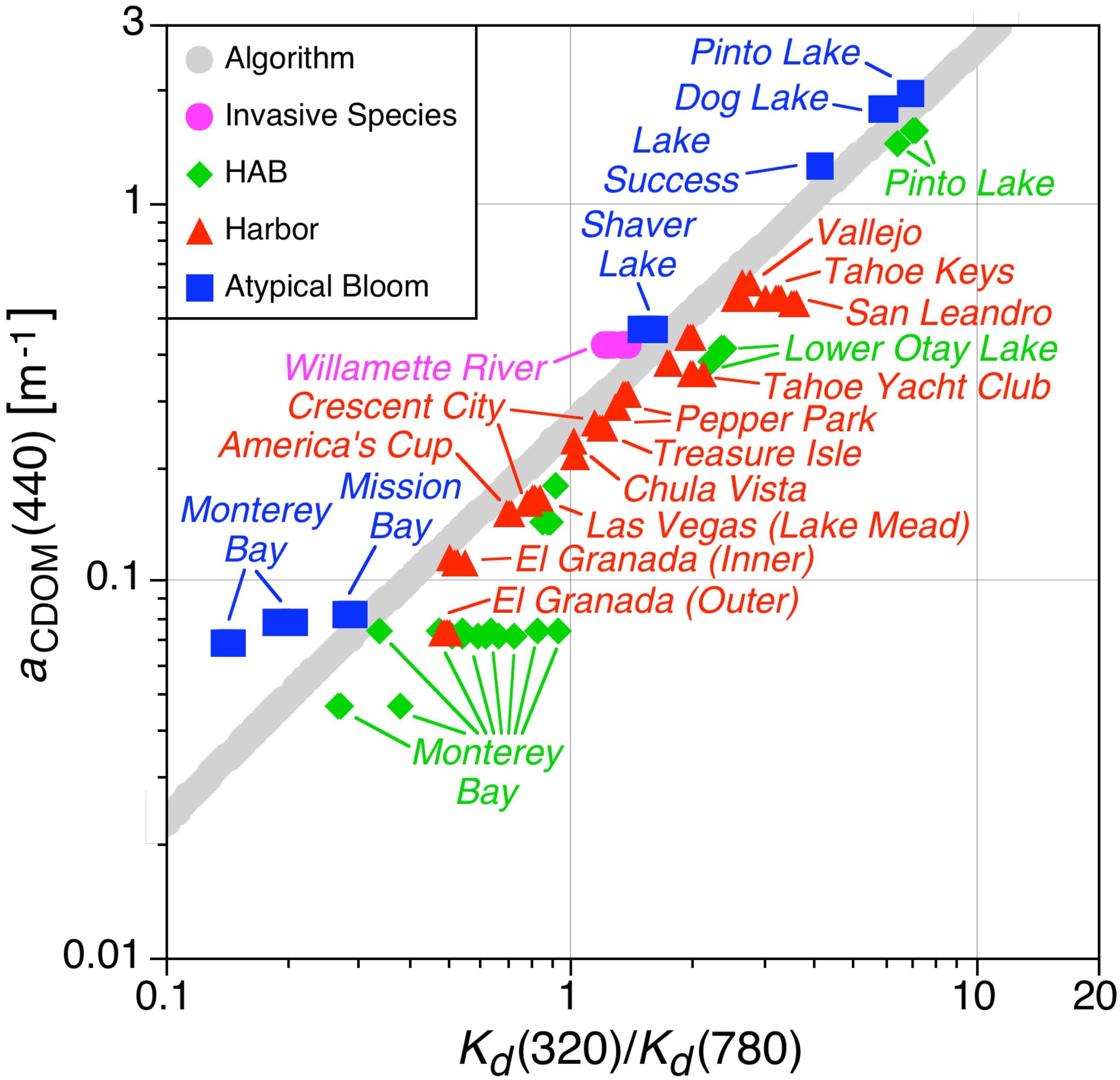
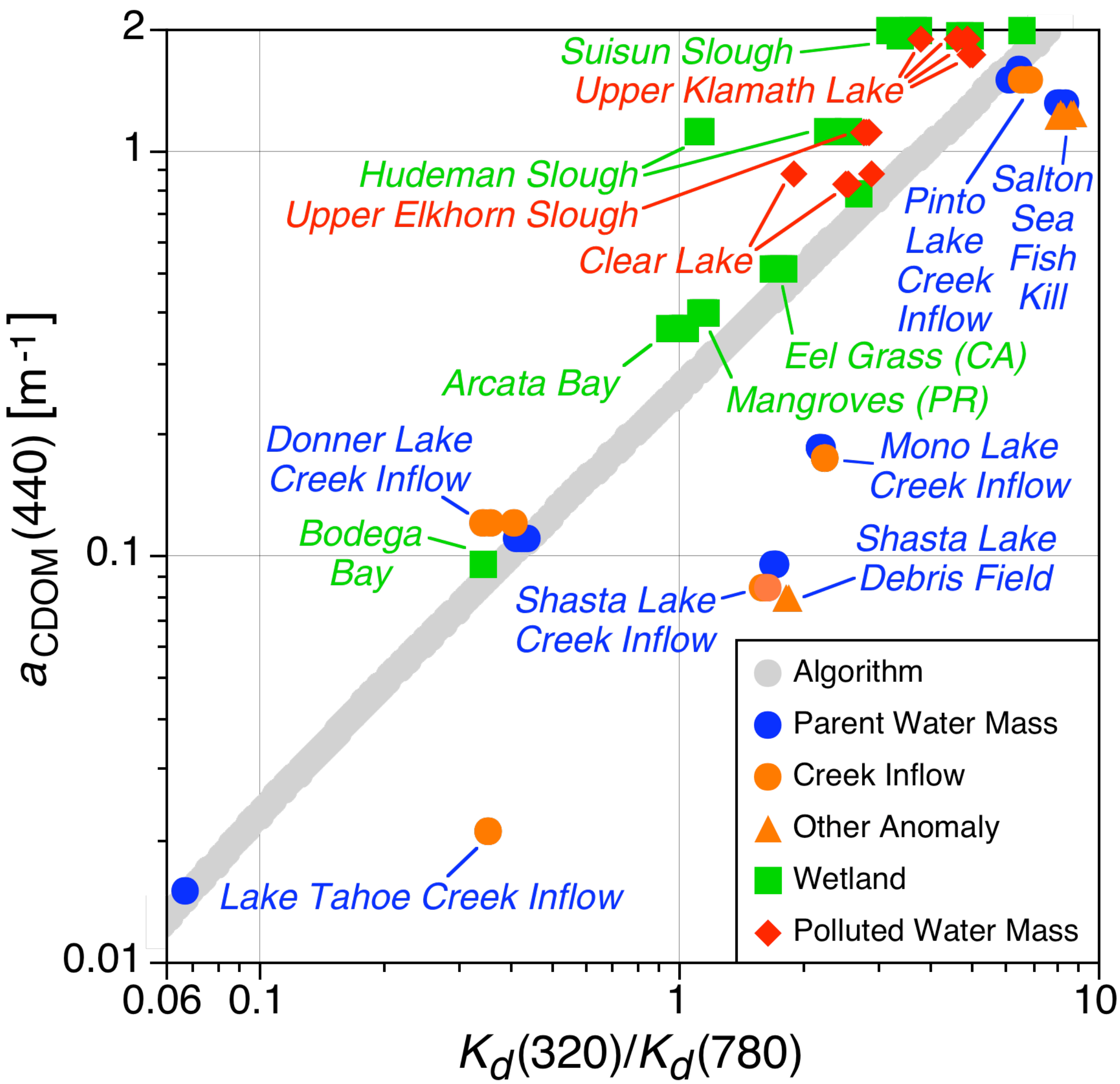
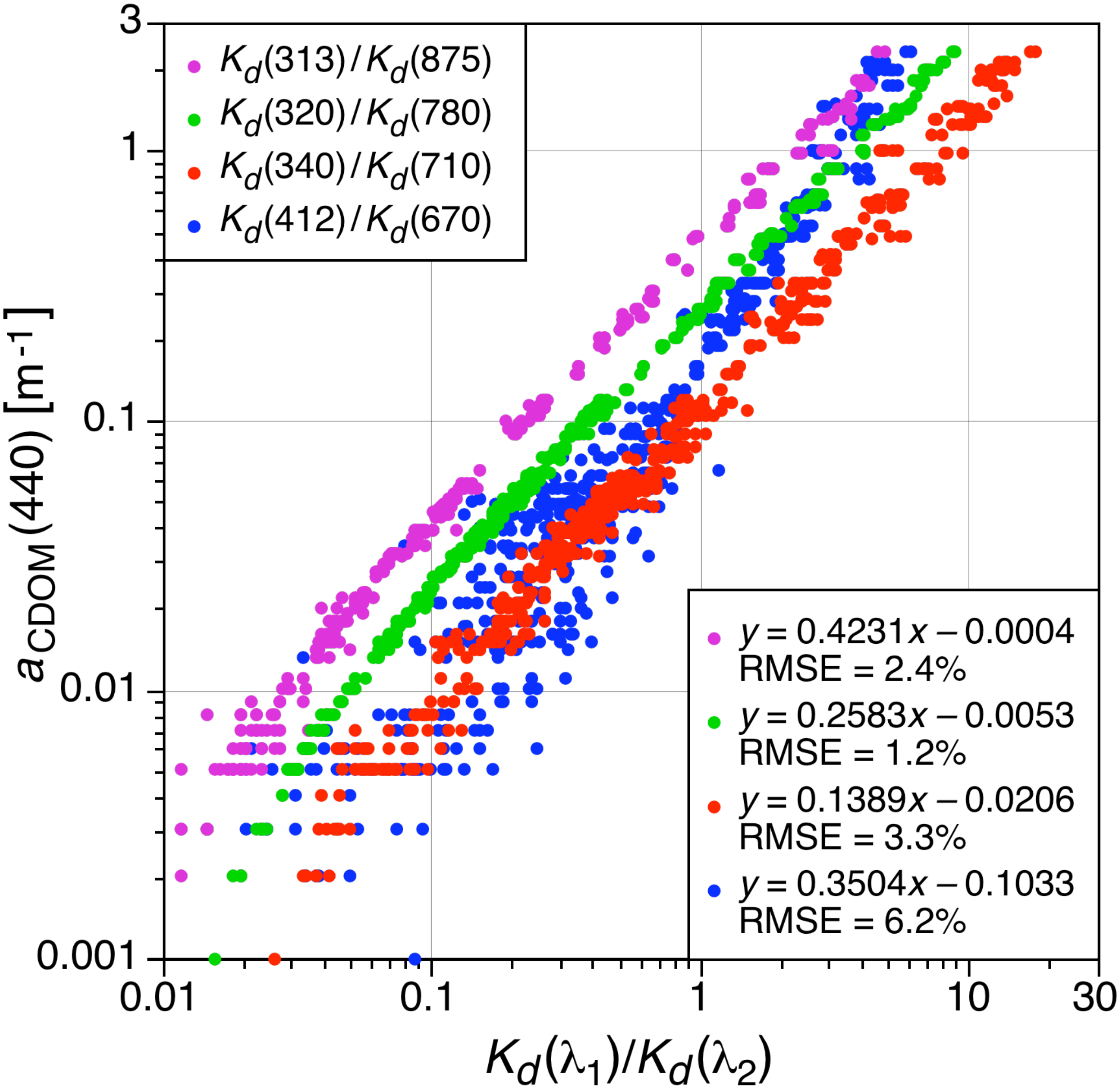

Algorithm
Invasive Species
HAB
Harbor
Atypical Bloom
Pinto Lake
Dog Lake
Lake Success
Pinto Lake
Shaver Lake
Vallejo
Tahoe Keys
San Leandro
Willamette River
Lower Otay Lake
Tahoe Yacht Club
Crescent City
Pepper Park
America's Cup
Treasure Isle
Mission Bay
Chula Vista
Monterey Bay
Las Vegas (Lake Mead)
El Granada (Inner)
El Granada (Outer)
Monterey Bay
3
1
0.1
0.01
0.1
1
10
20
$a_{\mathrm{CDOM}}(440)$ [m-1]
$K_d(320)/K_d(780)$

Figure 7



2
1
0.1
0.01
0.06
0.1
1
10
$a_{CDOM}(440)$ [m^{-1}]
$K_d(320)/K_d(780)$
Suisun Slough
Upper Klamath Lake
Hudeman Slough
Upper Elkhorn Slough
Clear Lake
Arcata Bay
Eel Grass (CA)
Mangroves (PR)
Pinto Lake Creek Inflow
Salton Sea Fish Kill
Donner Lake Creek Inflow
Bodega Bay
Mono Lake Creek Inflow
Shasta Lake Debris Field
Shasta Lake Creek Inflow
Lake Tahoe Creek Inflow
Algorithm
Parent Water Mass
Creek Inflow
Other Anomaly
Wetland
Polluted Water Mass

Figure 8



3
1
0.1
0.01
0.001
0.01
0.1
1
10
30
$a_{\mathrm{CDOM}}(440)$ [m-1]
$K_d(\lambda_1)/K_d(\lambda_2)$
$K_d(313)/K_d(875)$
$K_d(320)/K_d(780)$
$K_d(340)/K_d(710)$
$K_d(412)/K_d(670)$
$y = 0.4231x - 0.0004$
RMSE = 2.4%
$y = 0.2583x - 0.0053$
RMSE = 1.2%
$y = 0.1389x - 0.0206$
RMSE = 3.3%
$y = 0.3504x - 0.1033$
RMSE = 6.2%

FIGURE 9

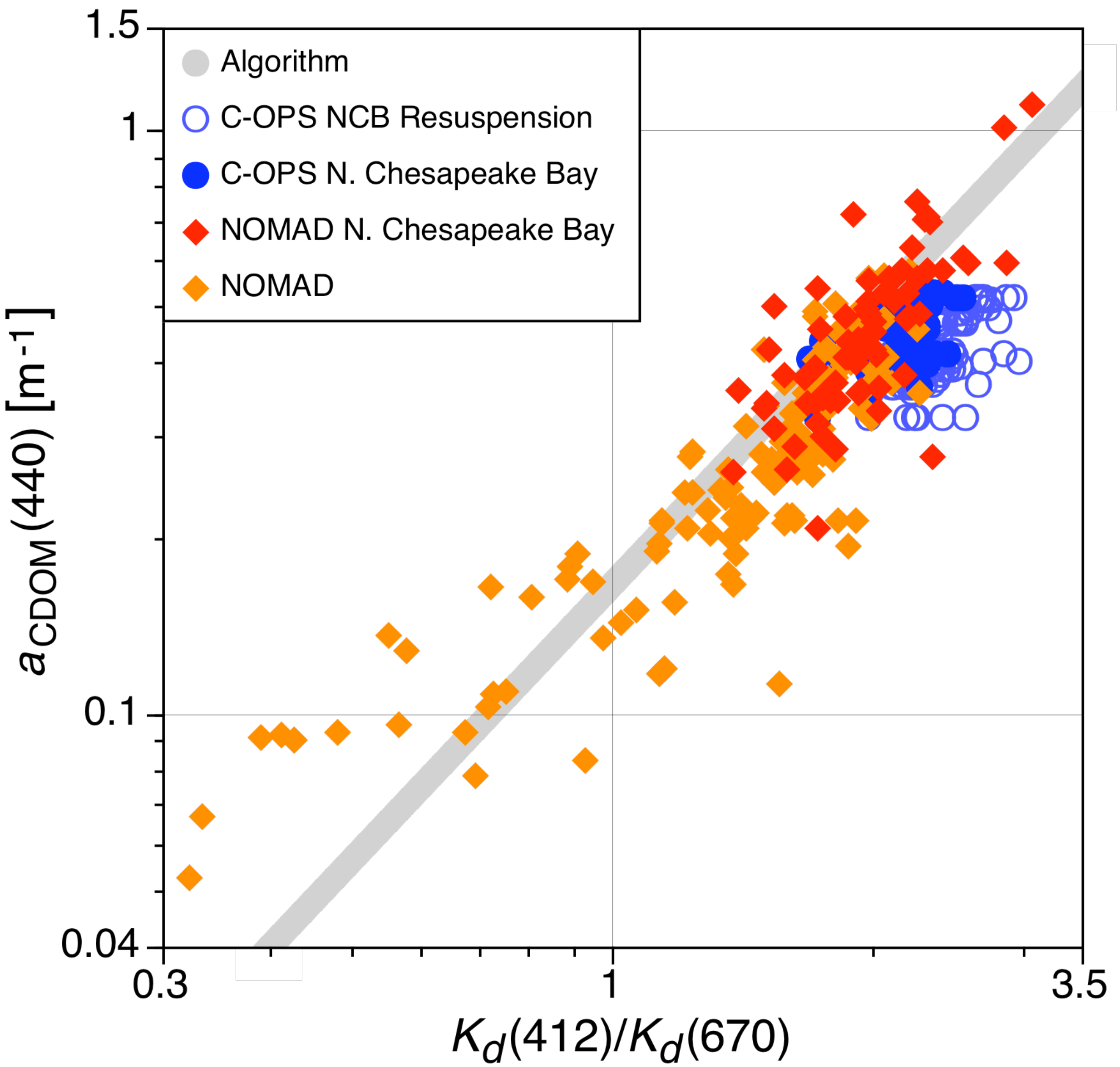
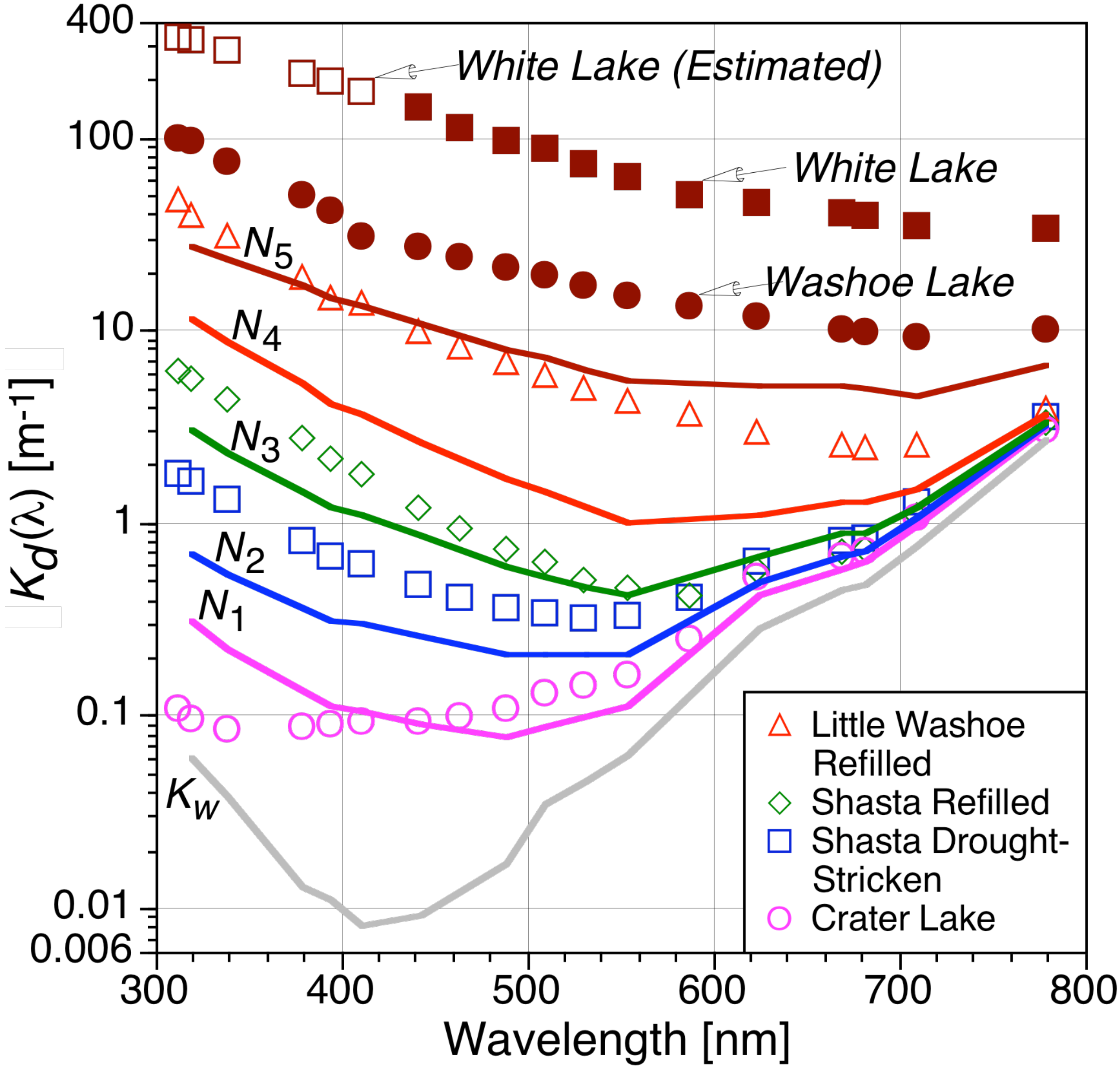

1.5
1
0.1
0.04
0.3
1
3.5
Algorithm
C-OPS NCB Resuspension
C-OPS N. Chesapeake Bay
NOMAD N. Chesapeake Bay
NOMAD
$a_{CDOM}(440)$ [m^{-1}]
$K_d(412)/K_d(670)$

Figure 10


White Lake (Estimated)
White Lake
Washoe Lake
N_5
N_4
N_3
N_2
N_1
K_w
Little Washoe Refilled
Shasta Refilled
Shasta Drought-Stricken
Crater Lake
400
100
10
1
0.1
0.01
0.006
$K_d(\lambda)$ [m^{-1}]
300
400
500
600
700
800
Wavelength [nm]