

EARTH SYSTEM SCIENCE DATA

Ref: ESSD-2025-438

Response letter to Reviewers' Comments

Reviewer #1

Comments to the Author

The manuscript aims to construct a high-quality R_{rs} dataset covering coastal waters of both China and the United States. An AC processing framework that integrates standard NIR-based AC and grided aerosol model AC were developed. Moreover, cloud masking constraint was relaxed to get more effective data. While the topic is important as R_{rs} products in coastal areas is still have relatively high uncertainty as compared to it in open oceans, the high quality promoted by the title does not match the actual results for the following four major reasons.

Response: Thank you for your thorough evaluation and the critical points raised, which allow us to better articulate the advances and rationale of our work. We address each concern below, demonstrating that our integrated framework achieves a superior balance of data quality, accuracy, and coverage specifically for the challenging coastal waters it targets.

1. Relaxing cloud masking constraint of course would enlarge effective data portion, but it would degrade data quality or accuracy of the R_{rs} products especially in clear waters. Therefore, this contradicts the goal of obtaining high-quality R_{rs} dataset.

Response: Thank you for raising this important point. We fully acknowledge that indiscriminate cloud mask relaxation would degrade data quality if applied universally. However, a key premise of our study is its focus on optically complex coastal and turbid waters, where the standard NASA product suffers from severe data loss due to overly conservative masking.

To address this data loss, we implemented a revised cloud mask threshold ($R_{rc,2130}=0.037$) not as a standalone change, but as the first step in a targeted QC framework designed explicitly for this challenging environment. This threshold, validated for coastal/inland waters by Zhang et al. (2014), maximizes the retention of valid turbid-water pixels that are often misclassified as cloud. Crucially, it is followed by a subsequent “Wrong Turbid Water Mask” filter that actively removes pixels with spectral signatures uncharacteristic of true turbid water (e.g., anomalously high blue-band reflectance), thereby safeguarding against the inclusion of thin clouds or other non-water features. Furthermore, to ensure the radiometric reliability of the retained water pixels themselves, our QC chain also applies a “Saturated $R_{rs}(667)$ Mask”. This

filter excludes pixels with saturated signals in the 667 nm band, which are indicative of extremely turbid waters where AC becomes highly uncertain (Hu et al., 2012).

Thus, the relaxation and the subsequent targeted filtering steps operate together within an integrated QC chain. This approach prioritizes recovering critical data coverage in our target domain (turbid coastal waters) while ensuring reliability by proactively removing both non-water features and water pixels with unreliable retrieval conditions. The effectiveness of this balance is confirmed by the high quality of the additionally retrieved pixels, which show physically plausible spatial patterns rather than the speckled noise typical of cloud contamination (e.g., Figure 8).

We acknowledge that cloud masking over dynamic global waters remains a challenging research topic. Should more sophisticated, regionally optimized masking algorithms emerge in the future, they could be integrated into our processing chain for further refinement. For the present study, focused on nearshore turbid waters, the implemented framework represents a robust and effective solution that significantly improves upon the status quo.

To clearly present this integrated design in the manuscript, we have revised the paragraph regarding cloud mask relaxation in Section 2.2.2 (“Quality Control Strategies”) as follows: **“Among these flags, the CLDICE mask was customized for turbid coastal waters by applying a revised threshold ($R_{rc,2130}=0.037$). This adjustment reduces excessive cloud masking over areas with high AOT and turbid waters (Zhao et al., 2023; Zhang et al., 2014) and serves as the targeted first step to recover valid turbid-water pixels that are frequently misclassified.”** We have also revised the concluding statement of this subsection as: **“By implementing this multi-step QC framework—comprising initial flagging, targeted cloud mask relaxation with subsequent spectral filtering, and exclusion of unreliable water signals—we prioritize data recovery in challenging coastal regions while proactively controlling for quality, thereby achieving a balance between coverage and reliability. Applying these QC strategies consistently across both datasets ensured that subsequent analyses were based on equivalently screened, quality-assured R_{rs} products.”**

Furthermore, we have added a new Section 4.3 (“Considerations on Quality Control Strategies”) to discuss the effect and design choices of our QC framework and outline future work. In this section, we reference comparative findings from the newly added Section 4.1, “Comparison Against Other AC Algorithms for Coastal Waters”, to contextualize the discussion (the details of Section 4.1 are provided later in this response letter). The revised Section 4.3 reads:

The QC strategy is integral to the reliability of any ocean color dataset. Our framework combines a relaxed cloud mask ($R_{rc,2130}=0.037$) with additional filters (e.g., the “Wrong Turbid Water Mask”), to prioritize the retention of valid turbid

water pixels that are routinely discarded by standard processors, while implementing safeguards against anomalous retrievals.

However, when comparing with OC-SMART and POLYMER (Section 4.1, Figure 14), we observed that their default masks retained slightly more valid pixels than our more conservative QC chain. This difference is primarily attributable to our “Wrong Turbid Water Mask,” (See Figure 8), which excluded some pixels based on spectral criteria. This illustrates an inherent tension in QC design for coastal waters: an overly conservative spectral filter may reject some valid but optically extreme water pixels, whereas a more permissive mask risks retaining subtle cloud contamination or other artifacts. Our current strategy prioritizes per-pixel confidence over initial data yield, accepting a minor reduction in the latter for enhanced overall reliability.

Looking forward, cloud and cloud-shadow detection remains one of the most persistent challenges in global aquatic remote sensing. Although the QC framework used in this study performs robustly in the target regions, its universal optimization is non-trivial due to the vast diversity of atmospheric and aquatic conditions. The development and integration of more advanced, physically based or machine-learning-driven cloud masking algorithms, capable of regional adaptation and distinguishing between high turbidity and true cloud cover, represents a critical future direction. Such advancements could potentially obviate the need for heuristic spectral filters like the “Wrong Turbid Water Mask”, further automating the pipeline and enhancing the consistency and yield of quality-controlled data products across all coastal zones.

2. The representativeness of gridded aerosol models in instantaneous satellite images of coastal or shelf waters is questionable. This issue includes two parts. One is the temporal representativeness, since the gridded aerosol models are multi-years and monthly averaged data, it can't represent the instantaneous aerosol conditions. For example, even for some coastal oceans that are often affected by absorbing aerosols, the aerosol type may still be non or weak-absorbing aerosols in a specific day's image. In this case, if the average absorbing aerosol is still used, the accuracy may decrease. Another one is the spatial representativeness, aerosol types are high spatial heterogeneity, the large grid size (5 degree) makes it difficult to ensure that the gridded aerosol models can represent the real aerosol types of each pixel, especially for grids with only one AERONET station far from the coastline.

Response: Thank you for your insightful comments on the spatiotemporal representativeness of the gridded aerosol models. These are important considerations.

On the temporal aspect, we fully acknowledge that monthly climatological models cannot perfectly represent the instantaneous aerosol state. However, in coastal regions where aerosol constituents have strong geographical or seasonal patterns, they provide a statistically more realistic prior than the default AF10 models, which are biased

toward open-ocean, non-absorbing aerosols. Critically, this prior is dynamically integrated into the standard NIR iterative process. It provides an optimized starting point, and the aerosol contribution is then refined based on the real-time observed signal. This design allows the retrieval to converge toward a correct solution even if the prior is not perfect, as long as it is within a reasonable range. In contrast, the standard algorithm's model selection, based on instantaneous spectral matching (ϵ), carries a different risk: when no candidate model matches the observation, it defaults to a non-absorbing aerosol model (e.g., r75f01), virtually guaranteeing error in the presence of absorbing aerosols. Our climatology-informed prior avoids this “default selection failure”, providing a more stable and physically plausible starting point for AC.

In terms of spatial resolution, we agree that a 5° grid may not capture very fine-scale aerosol gradients. The spatial resolution of our current models is primarily constrained by the global distribution of AERONET sites, whose high-quality, ground-based measurements are crucial for ensuring the optical accuracy of the aerosol models. This represents a deliberate trade-off, prioritizing model accuracy over spatial resolution at this stage of development. We note that for our study regions along the coasts of China and the US, the density of AERONET sites is relatively high, supporting the applicability of the 5° climatology for characterizing the dominant aerosol regimes within these regions.

Most importantly, the comprehensive validation against in-situ measurements (Figure 3) demonstrates that our product significantly reduces the systematic bias present in the standard NASA product (regression slope improved from 0.90 ± 0.06 to 1.00 ± 0.08), while maintaining comparable precision. This result confirms that, despite the acknowledged limitations of gridded models at fine spatiotemporal scales, our integrated approach—which uses them as an informed prior within an iterative scheme—yields a substantial improvement in retrieval accuracy for the target coastal waters.

We recognize enhancing spatiotemporal resolution is a key future goal, and we have begun exploring the integration of data from next-generation satellite sensors (e.g., polarimetric measurements from PACE) to develop higher-resolution aerosol models in the future.

In response to your insightful comment, we have added Section 4.4 (“Considerations on Model Representativity”) into the Discussion Section of the revised manuscript:

The performance of our AC framework is inherently linked to the quality of the aerosol prior. The representational characteristics of the gridded aerosol models—namely, their spatial (5°) and temporal (monthly) resolution—thus constitute a fundamental aspect of the framework’s design. While these models have proven effective, their fixed resolution may not fully capture the high heterogeneity of aerosols in coastal areas at instantaneous and pixel scales. However, as integrated into our processing chain (Section 2.2.1), the model serves

as an optimized starting point for iterative refinement, a design that mitigates the impact of imperfect prior matching. The validation results (Figure 3) and scene-based comparison (e.g., Figure 8) confirm that this approach yields a significant and effective correction of the systematic bias at the current model resolution. Future work to further enhance fine-scale coastal retrievals could therefore focus on developing regional aerosol climatology or integrating real-time assimilation products with higher spatiotemporal resolution, thereby providing an even more accurate AC starting point.

3. There is no scientific basis for using water turbidity to distinguish between gridded aerosol models and traditional AF10 aerosol models. The regulation mechanisms of water turbidity and aerosols are different, and there is no necessary connection between the two.

Response: Thank you for this comment, which allows us to clarify the rationale behind our spatially adaptive aerosol model selection strategy.

We agree that there is no direct physical mechanism linking water turbidity to aerosol characteristics. Our approach is a pragmatic strategy leveraging a strong spatial co-occurrence driven by proximity to the coast. Empirically, in nearshore regions strongly influenced by land-based processes, we typically observe both turbid waters (from river discharge and sediment resuspension) and a higher prevalence of continentally influenced, often absorbing aerosols. Conversely, farther offshore, clearer waters and maritime aerosols are commonly found together. While not causally linked, this consistent spatial pattern allows water turbidity, as measured by the normalized water-leaving radiance in the NIR band, to serve as a robust practical proxy for estimating the likelihood of encountering continental versus maritime aerosols at a given pixel. Therefore, it provides an effective criterion for dynamically selecting the most appropriate aerosol model in a given pixel.

In coastal areas, adopting a fixed aerosol model suite across such strong land-ocean gradient is problematic: using gridded models offshore can overestimate aerosol absorption in clear-water regions, while persisting with AF10 models nearshore perpetuates known underestimation biases. Our dynamic switching strategy is therefore necessary to mitigate these opposing errors and generate a spatially consistent R_{rs} product.

The effectiveness of this pragmatic approach is demonstrated by the results. The smooth spatial gradients in the resulting R_{rs} imagery (e.g., Figure 8, 9), without discontinuities that would arise from inappropriate or abrupt model switches, provide visual evidence that our proxy-based selection strategy functions as intended across the land-ocean gradient. The comprehensive validation (Figure 3) also shows a significant reduction in systematic bias for the final product.

In response to your comment regarding the basis for using water turbidity, we have expanded the discussion in the manuscript to provide a clearer and more nuanced

rationale. A new subsection (4.2 “Considerations on the Aerosol Model Selection Strategy”) has been added within the Discussion Section:

A pivotal component of our improved AC framework is the dynamic selection of aerosol models, guided by the normalized water-leaving radiance in the NIR band ($[L_w(NIR)]_N$). This design is predicated on the strong empirical spatial co-occurrence observed in coastal zones: waters exhibiting high turbidity (and thus elevated $[L_w(NIR)]_N$) are typically nearshore, and this high turbidity spatially co-occurs with prevalent continentally influenced, often absorbing aerosols. Conversely, in offshore regions, waters are generally clear (with low $[L_w(NIR)]_N$) and maritime aerosols prevail. We leveraged this consistent geographic pattern as a practical and effective proxy to implement a necessary transition between aerosol models across the land-ocean gradient. This spatially adaptive selection prevents the systematic underestimation that would occur if AF10 models were applied to turbid nearshore waters, while simultaneously avoiding the overestimation that would arise if gridded models representing absorbing aerosols were used in clear offshore waters. The success of this approach is evidenced by the reduction of systematic bias (Figure 3) and the spatially coherent R_{rs} fields without discontinuities (e.g., Figures 8, 9).

It is important to clarify that this strategy does not assert a direct physical causation between water optical properties and aerosol characteristics. Rather, it is a pragmatic solution that capitalizes on their correlated spatial distributions driven by common terrestrial influences. Consequently, the performance of this approach is inherently tied to the robustness of this co-occurrence relationship. In regions or under conditions where this relationship weakens or decouples (e.g., during long-range transport of dust or smoke over clear ocean), the proxy may be less effective. To overcome this limitation and extend the framework's applicability, research could focus on integrating direct, physically based aerosol indicators (e.g., from multi-angle or polarimetric sensors). Alternatively, developing regionally adapted or higher-resolution co-occurrence models could refine the proxy and enhance the robustness of the aerosol selection process.

4. The standard NIR AC algorithm has a high degree of uncertainty in turbid water, which has been widely recognized. Therefore, comparing the results with NIR AC algorithm in this study cannot fully demonstrate the superiority of the proposed method. It is recommended to compare it with existing atmospheric correction results for turbid water, such as OC-SMART, POLYMER, NIR-SWIR, etc.

Response: Thank you for your constructive suggestion on providing a comparison with additional state-of-the-art AC algorithms designed for turbid waters to assess our framework more comprehensively.

We have addressed this by adding a detailed comparative analysis into the manuscript as a new subsection (Section 4.1). This analysis directly compares our improved

framework against three advanced algorithms (OC-SMART, POLYMER, NIR-SWIR) using four representative turbid water scenes. The results show that our framework retrieves the highest R_{rs} values in the most turbid regions, providing the most effective correction of the systematic low bias present in the standard product. In contrast, the corrective effect of the other algorithms in these extreme conditions appears limited, as their outputs are frequently comparable to a baseline that retains the known low bias (AF10&2130). The discussion also extends to the methodological implications, noting that our modular, physics-based design offers a more straightforward path for regional adaptation compared to strategies reliant on pre-trained global spectral models. The revised text added to the manuscript is as follow:

4.1 Comparison Against Other AC Algorithms for Coastal Waters

To further contextualize the performance of our framework, we conducted a scene-based intercomparison with three advanced AC algorithms for complex waters: OC-SMART(Fan et al., 2021), POLYMER(Steinmetz et al., 2011), and NIR-SWIR, using the same four representative regions from Section 3.3. This analysis also included results from the standard NIR iterative algorithm with AF10 aerosol models and relaxed cloud mask (AF10&2130, as mentioned before). This configuration provides a reference baseline that embodies the persistent low bias, so the corrective effect of other algorithms can be assessed against it. For a fair and controlled comparison, the NIR-SWIR and AF10&2130 configurations used the same cloud mask and quality control steps as our new dataset, while OC-SMART and POLYMER were run with their default settings.

The comparative imagery (Figures 14-17) demonstrates the superior performance of our framework. While all five methods successfully capture major spatial patterns associated with turbid water features (e.g., sediment plumes) in the 547 nm and 667 nm bands, our new dataset consistently retrieves the highest R_{rs} values in the most turbid regions across all scenes and bands (443, 547, 667 nm). In contrast, the R_{rs} values retrieved by OC-SMART, POLYMER, and NIR-SWIR in these challenging areas are frequently comparable to, or only marginally higher than, those of the AF10&2130 configuration. Given that the AF10&2130 configuration is quantitatively demonstrated to suffer from systematic underestimation (Section 3.1, Figure 3), this indicates that the corrective effect of these other algorithms on the low bias is limited in optically extreme waters, whereas our improved framework provides the most robust correction.

Regarding data quality, NIR-SWIR results exhibit increased image noise due to the low signal-to-noise ratio of MODIS Aqua's SWIR bands. While POLYMER and OC-SMART generally perform well, isolated issues such as unmasked anomalous pixels (Bohai Sea for POLYMER) and striping artifacts (Mississippi River Estuary at 443 nm for OC-SMART) were observed. In terms of data coverage, the default masks for OC-SMART and POLYMER retained slightly more data in the Pearl River Estuary than our more conservative QC chain.

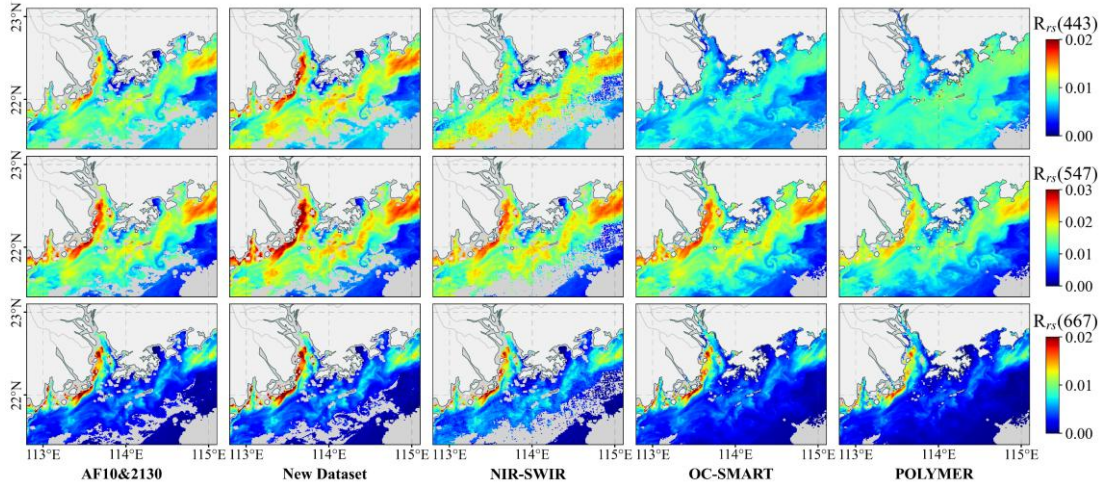


Figure 14: Comparison of R_{rs} (units: sr^{-1}) derived from single-scene MODIS Aqua observations over the Pearl River Estuary (2014 12-14). The figure shows R_{rs} at 443 nm (top row), 547 nm (middle row), and 667 nm (bottom row) as processed by the AF10&2130 configuration (first column), our new dataset (second column), the NIR-SWIR algorithm with AF10 aerosol models and $R_{rc,2130}=0.037$ as cloud mask threshold (third column), OC-SMART (fourth column), and POLYMER (fifth column).

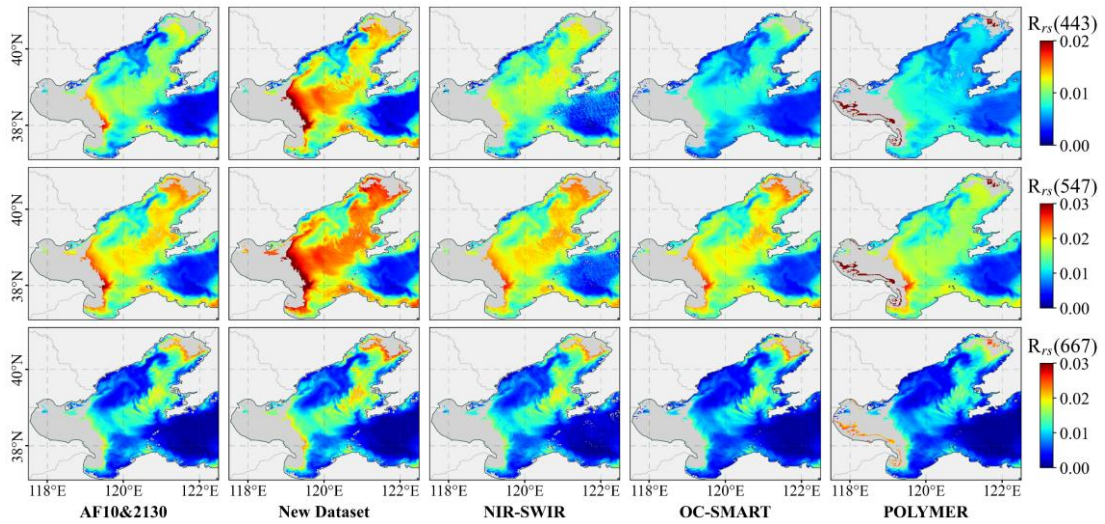


Figure 15: Comparison of R_{rs} (units: sr^{-1}) derived from single-scene MODIS Aqua observations over the Bohai Sea (2020-03-11). The figure shows R_{rs} at 443 nm (top row), 547 nm (middle row), and 667 nm (bottom row) as processed by the AF10&2130 configuration (first column), our new dataset (second column), the NIR-SWIR algorithm with AF10 aerosol models and $R_{rc,2130}=0.037$ as cloud mask threshold (third column), OC-SMART (fourth column), and POLYMER (fifth column).

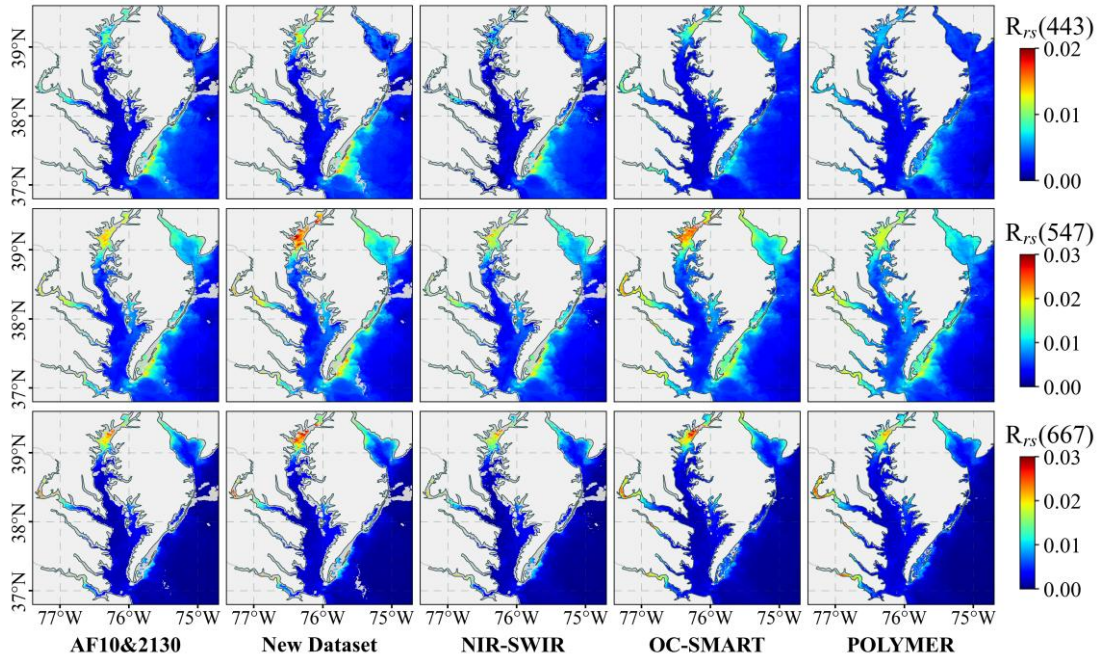


Figure 16: Comparison of R_{rs} (units: sr^{-1}) derived from single-scene MODIS Aqua observations over the Chesapeake Bay (2019-02-05). The figure shows R_{rs} at 443 nm (top row), 547 nm (middle row), and 667 nm (bottom row) as processed by the AF10&2130 configuration (first column), our new dataset (second column), the NIR-SWIR algorithm with AF10 aerosol models and $R_{rc,2130}=0.037$ as cloud mask threshold (third column), OC-SMART (fourth column), and POLYMER (fifth column).

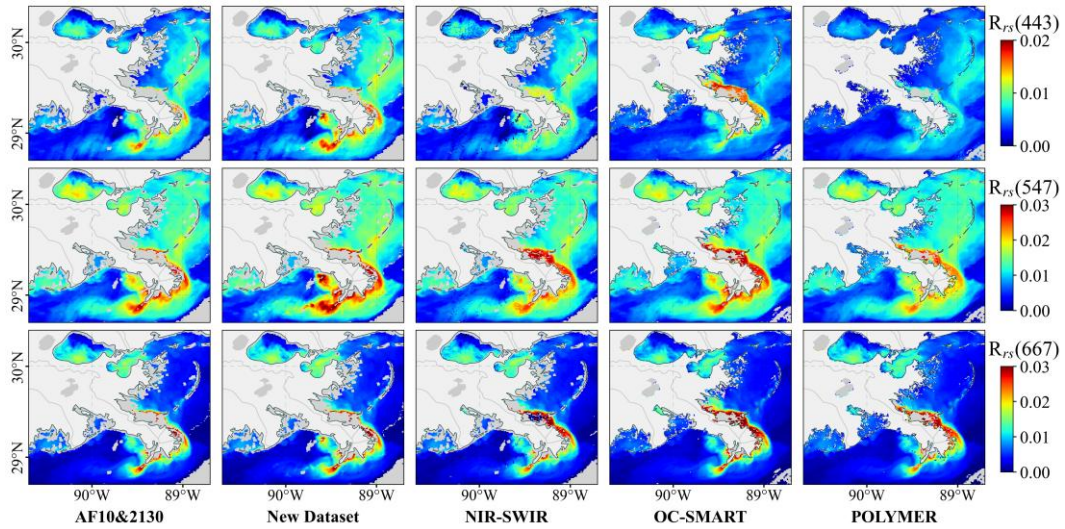


Figure 17: Comparison of R_{rs} (units: sr^{-1}) derived from single-scene MODIS Aqua observations over the Mississippi River Estuary (2021-01-13). The figure shows R_{rs} at 443 nm (top row), 547 nm (middle row), and 667 nm (bottom row) as processed by the AF10&2130 configuration (first column), our new dataset (second column), the NIR-SWIR algorithm with AF10 aerosol models and $R_{rc,2130}=0.037$ as cloud mask threshold (third column), OC-SMART (fourth column), and POLYMER (fifth column).

Beyond these empirical results, our methodological approach itself warrants discussion. It enhances a physically based, modular component (aerosol modeling) within the radiative transfer framework. This differs from strategies that depend

on comprehensively pre-trained spectral or bio-optical models (e.g., POLYMER, OC-SMART), whose performance is inherently tied to the optical domains covered during their training or optimization. Adapting our framework to a new coastal region would primarily involve updating the modular aerosol prior with local or regional climatological data. In contrast, methods like POLYMER or OC-SMART may require significant recalibration or retraining to maintain performance outside their original optimization domain. Thus, our design offers a more straightforward and transparent path for regional scalability.

We believe these additions address your concern and enhance the depth and completeness of the “Data and Method” and “Discussion” sections. Thank you again for your insightful comment, which has helped improve the quality of our manuscript.

Reviewer #2

Comments to the Author

This manuscript describes a dataset of “high-quality” remote sensing reflectance (R_{rs}) over the coastal waters off China and the United States. R_{rs} is an important water property; thus, this manuscript is appropriate for this journal. In addition, the manuscript is well written, very easy to follow. However, this reviewer does not think it is ready for publication, at least for the present version.

Response: Thanks for your comments. We have carefully revised this manuscript based on the reviewer’s comments. We hope that the revisions and improvements would satisfactorily address the reviewer’s concerns. Our point-to-point response is as follows.

Main issue:

1. The R_{rs} product of the global ocean, including coastal waters, is already distributed by NASA (and many other agencies) for decades. Publication and distribution of newly generated products should demonstrate an overall better quality compared to existing ones. However, as presented in Fig. 3, for the 8 wavelengths compared, NASA products show clearly better (in R^2 , MdAPE, NRMSE) quality for 4 wavelengths (412, 443, 667, and 678 nm), with the quality similar for the other 4 wavelengths (488, 531, 547, and 555 nm). These underperformances are also observed by the authors. In view of these, it would cause confusion to the user community if such a data product were distributed.

Response: Thank you for your critical and insightful comment. We agree that any new data product must demonstrate clear added value over established standards like the NASA ocean color products. However, we argue that the value of our dataset lies not in outperforming the standard product on every metric at every wavelength, but in offering a superior solution for the specific challenge of monitoring optically complex coastal waters, where the standard product has known limitations.

The reviewer correctly notes that for certain bands, our precision metrics (R^2 , MdAPE, NRMSE) are comparable to or slightly lower than the standard product’s. This is an understood consequence for achieving our primary objective: to correct the pervasive systematic low bias in NASA’s coastal retrievals. While the use of regionally optimized climatological prior inherently introduces some scene-specific variability (explaining the slightly lower precision metrics in certain bands), it successfully shifts the regression slope from 0.90 ± 0.06 to an ideal 1.00 ± 0.08 . From a practical standpoint, this trade-off is highly beneficial. For downstream biogeochemical applications (e.g., chlorophyll-a or SPM retrieval), eliminating systematic bias is more critical than minimizing random noise. This is because random variability can be effectively mitigated through spatial or temporal averaging, but a systematic bias will persist and propagate through all subsequent calculations. The slightly lower precision in specific blue and red bands is thus a localized cost of this bias correction, reflecting the extreme sensitivity of these spectral regions to aerosol and signal-to-noise challenges in coastal

zones. When weighed against the substantial improvement in systematic accuracy and the dramatic increase in data coverage (discussed below), this does not diminish the overall superior quality and utility of our dataset for coastal research.

Beyond radiometric metrics, a decisive advantage of our dataset is its significantly higher data coverage. In optically complex coastal zones, NASA's standard products often suffer from severe data loss due to overly conservative quality masks. In contrast, our dataset recovers substantially more valid data (e.g., +56% at 443 nm in Chinese coastal waters, Figure 4), thereby offering a more temporally and spatially complete observational record. This is critical for reliably monitoring short-term ecological events or resolving long-term trends—applications where the standard product's frequent data gaps pose a fundamental limitation.

Consequently, we explicitly position our dataset as a solution optimized for coastal and turbid water research, complementing rather than replacing the standard product. For the open ocean, the NASA product remains an excellent choice. For studies focused on estuaries, bays, and turbid coastal regions, however, our dataset provides a more robust, bias-corrected, and spatially more complete alternative.

To clearly define these application boundaries, we have added the following sentence into the manuscript's abstract: **“As a complementary resource optimized for estuaries, bays, and turbid coastal regions, the new dataset fills a critical gap in the ocean color data portfolio where standard products are less reliable.”** The Conclusion Section has also been correspondingly strengthened to reflect this specialized focus. Furthermore, we have revised Section 3.1 (“*In Situ* Validation”) of the manuscript in response to your comment on the presentation of validation results. The revisions sharpen the narrative to unequivocally highlight the correction of the pervasive systematic low bias as the primary achievement. For the blue and red bands where performance is more complex, the text now provides a clearer, objective interpretation that acknowledges inherent retrieval challenges and residual uncertainties, while demonstrating that the overall systematic low bias has nonetheless been significantly mitigated. The summary of the section is tightened to explicitly frame the advance as one achieved in “coastal” R_{rs} accuracy, thereby focusing the reader's assessment on the intended application domain.

In summary, by achieving a superior balance between bias correction, precision, and data availability, our dataset delivers “overall better quality” for coastal water studies, complementing the existing ocean color portfolio.

2. Fig. 8 shows more data products by the new approach. However, it is not clear if it is due to the new atmospheric correction algorithm, or due to the use of updated masks. If the NASA products were also processed by the updated mask, would the number of valid products be similar?

Response: Thank you for this insightful question. To directly address it, we have thoroughly revised Section 3.3 (“Single-Scene Imagery Comparison”) of the

manuscript to incorporate the “AF10&2130” configuration—applying the standard AF10 aerosol models with our updated cloud mask ($R_{rc,2130}=0.037$)—as a critical baseline within the comparative figures (updated Figures 8-11). The analysis clearly delineates that using the relaxed cloud mask could increase data coverage, but for achieving accurate R_{rs} values, integration of improved aerosol models is essential. The revised section is provided below for your review:

3.3 Single-Scene Imagery Comparison

To further assess the utility of the new R_{rs} dataset, we conducted single-scene comparisons across four representative coastal regions: the Pearl River Estuary, the Bohai Sea, the Chesapeake Bay, and the Mississippi River Estuary (see Figure 1 for their locations). This analysis aimed to evaluate the dataset’s ability to capture regional R_{rs} features under diverse optical conditions, and to delineate the respective contributions of the updated cloud mask and improved aerosol model. To this end, the comparisons include not only our new dataset and the standard product, but also the AF10&2130 configuration (defined in Section 3.2) that applies the standard AF10 aerosol models with our updated cloud mask ($R_{rc,2130}=0.037$).

Figure 8 and Figure 9 present the comparative results for the Pearl River Estuary and the Bohai Sea, both of which are characterized by turbid nearshore waters (Liu et al., 2012; Xu et al., 2018). In these regions, the new dataset provides a markedly more extensive coverage of valid data than the standard product. For instance, while the standard products yield few valid pixels within the Pearl River Estuary and large portions of the Bohai Sea, the new dataset successfully recovers these missing observations. This enables the capture of fine-scale coastal structures—such as the distinct high-reflectance band along the Pearl River Estuary’s western coast—that are otherwise lost. Critically, the AF10&2130 configuration achieves near-identical spatial coverage of valid pixels, confirming that the relaxed cloud mask is the primary driver for the increased data availability in our new dataset.

However, among the restored turbid water pixels, critical differences exist in the retrieved R_{rs} values. In specific regions characterized by intense sediment plumes, such as the western coast of the Pearl River Estuary and the central Bohai Sea, the R_{rs} values retrieved by the AF10&2130 configuration are systematically and significantly lower than those in our new dataset. This pattern is consistent with the known limitation of the standard AF10 models, which tend to overestimate aerosol path radiance in the presence of absorbing aerosols, thereby over-correcting the water-leaving signal and leading to a pervasive low bias. The regionally informed aerosol models integrated into our framework effectively mitigate this issue. Previously validated for strongly absorbing coastal regions like the Bohai Sea (Lv et al., 2024; Zhao et al., 2023), these models provide a more accurate representation of local aerosol properties. Consequently, within the

sample windows (marked by black rectangles in Figures 8 and 9), our new dataset yields higher and physically more plausible R_{rs} values. For example, the mean $R_{rs}(547)$ in the Pearl River Estuary window increases from approximately 0.0233 sr^{-1} (AF10&2130) to 0.0270 sr^{-1} , representing a more realistic spectral magnitude for these turbid waters.

We further examined two representative regions in the US coastal waters: the Chesapeake Bay and the Mississippi River Estuary. The same three-way comparisons presented in Figures 10 and 11 (new dataset, standard product, AF10&2130) corroborate the findings from the Chinese seas. The relaxed cloud mask remains the primary driver for increased data coverage, recovering valid pixels in areas (e.g., the northern Chesapeake Bay and western Mississippi River Delta) that are typically flagged in the standard product. Meanwhile, the improved aerosol model ensures the radiometric fidelity of these recovered pixels. Within the highlighted sample windows, the new dataset yields R_{rs} values that are consistently higher than the AF10&2130 baseline—for example, approximately 0.0242 sr^{-1} versus 0.0199 sr^{-1} in the Chesapeake Bay at 547 nm—further evidencing the mitigation of systematic underestimation. These consistent results across geographically and optically diverse waters confirm that the observed improvements are not region-specific.

The analysis of the QC masks provides further insight. The missing pixels in the standard product are primarily flagged as cloud- or ice-contaminated, which is a misclassification attributable to the high reflectance of turbid water at 869 nm. Our updated cloud mask successfully recognizes these pixels as valid water observations. Furthermore, the new dataset's complementary QC filters (e.g., the “Wrong Turbid Water Mask” and “Saturated $R_{rs}(667)$ Mask” introduced in Section 2.2.2) work in tandem with the relaxed cloud mask. These filters enhance overall reliability by screening out pixels with anomalous spectral features, ensuring that data quality is maintained even if the detection constraints are relaxed.

Nevertheless, the new dataset still fails to retrieve valid data in certain extremely turbid waters, such as portions of the western Bohai Sea, indicating remaining challenges in AC for waters with exceptionally high reflectance. In summary, the integrated framework underlying our new dataset delivers a substantial advancement: it significantly increases data availability in coastal zones through a tailored cloud mask, while concurrently ensuring the radiometric accuracy of the recovered data through optimized aerosol correction and specialized QC filters. This dual improvement enhances the dataset's suitability for detailed ocean color analysis and reliable water quality monitoring in optically complex nearshore environments.

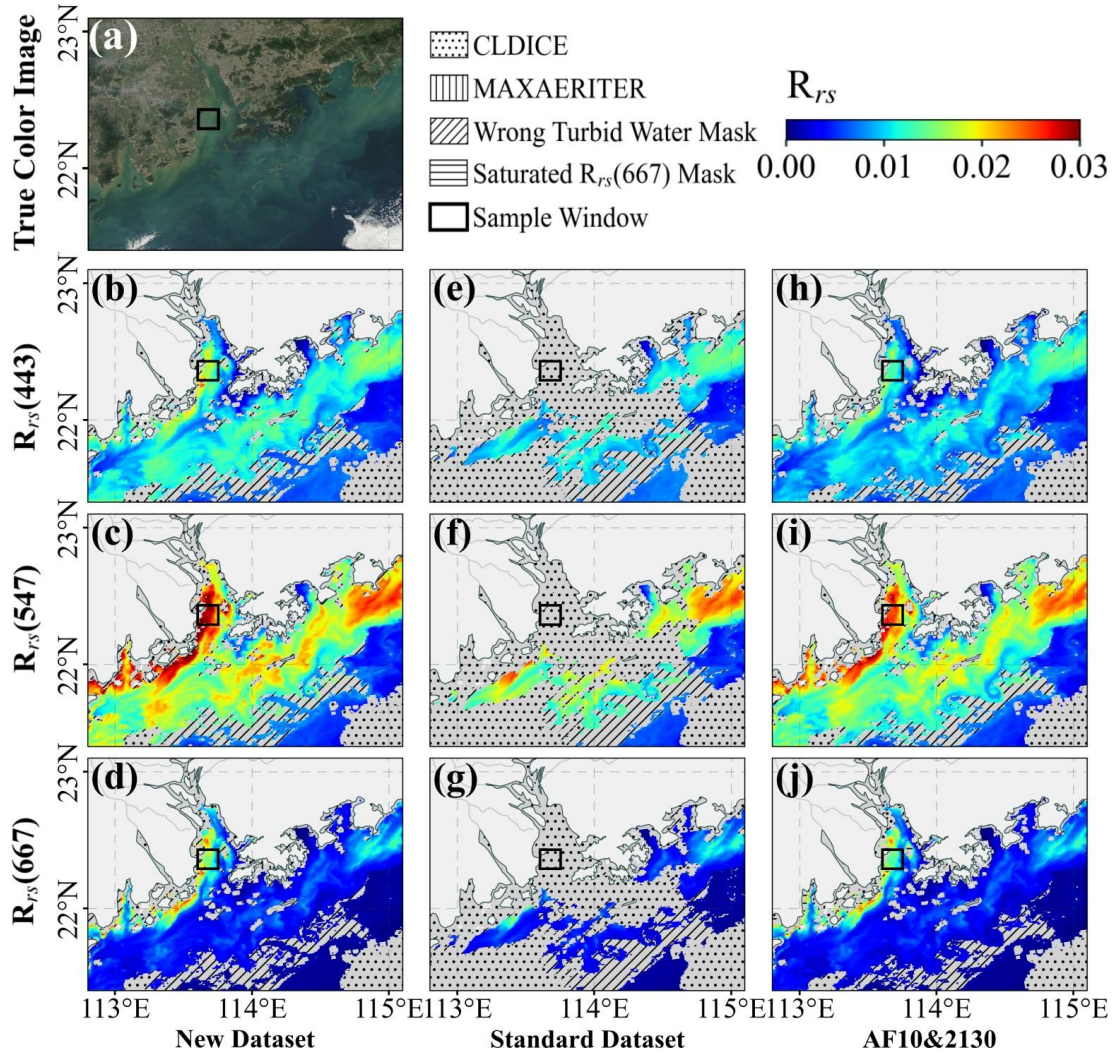


Figure 8: Comparison of R_{rs} (units: sr^{-1}) derived from single-scene MODIS Aqua observations over the Pearl River Estuary (2014-12-14). The top image shows a true color composite (panel a), followed by R_{rs} at 443, 547, and 667 nm from the new dataset (panels b–d), from the standard dataset (panels e–g) and from the AF10&2130 configuration (panels h–j). Invalid pixels are marked with hatched patterns based on quality control masks. Black rectangles mark the sample windows where average R_{rs} values are calculated and compared quantitatively in the text.

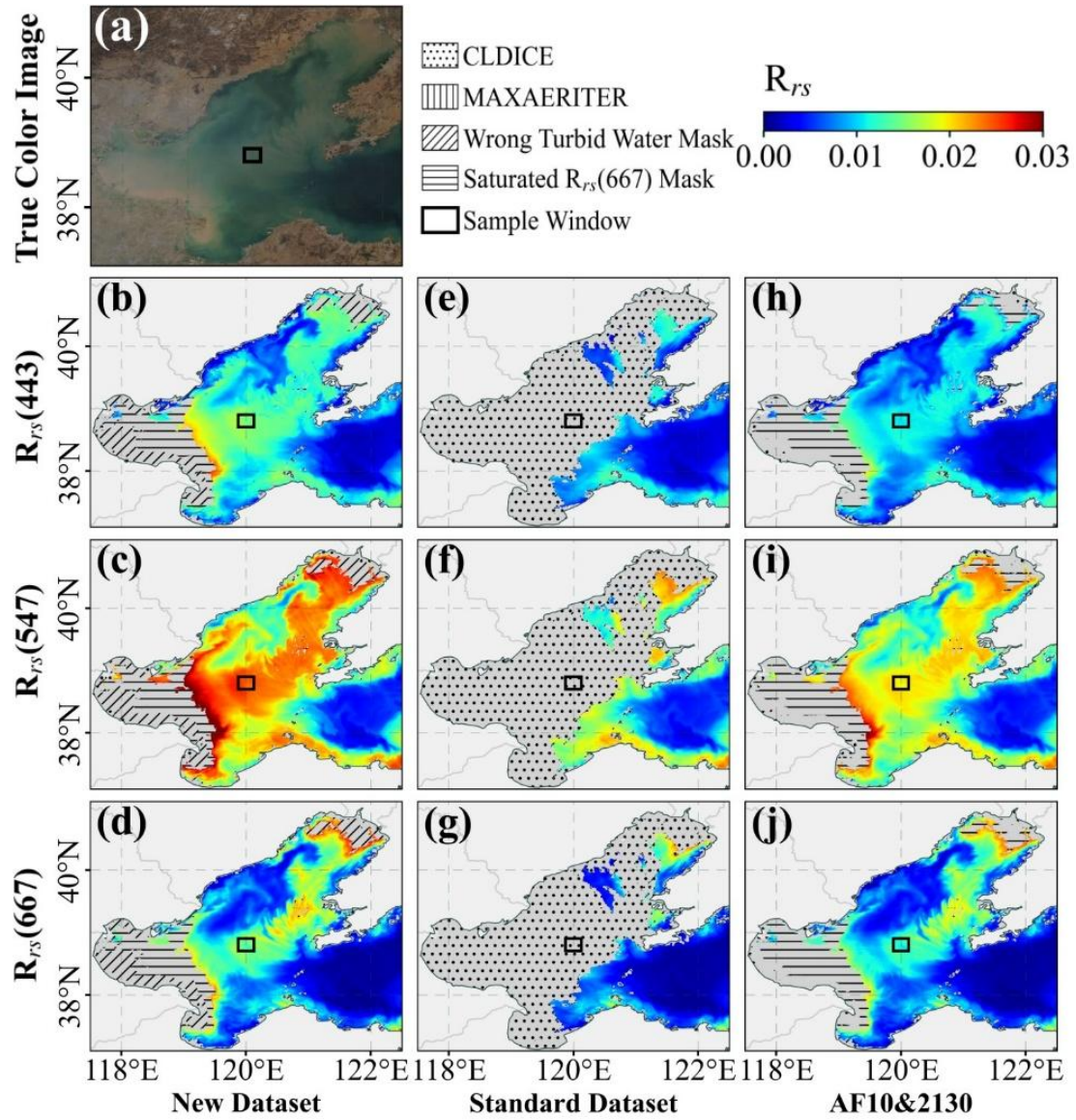


Figure 9: Comparison of R_{rs} (units: sr^{-1}) derived from single-scene MODIS Aqua observations over the Bohai Sea (2020-03-11). The top image shows a true color composite (panel a), followed by R_{rs} at 443, 547, and 667 nm from the new dataset (panels b–d), from the standard dataset (panels e–g) and from the AF10&2130 configuration (panels h–j). Invalid pixels are marked with hatched patterns based on quality control masks. Black rectangles mark the sample windows where average R_{rs} values are calculated and compared quantitatively in the text.

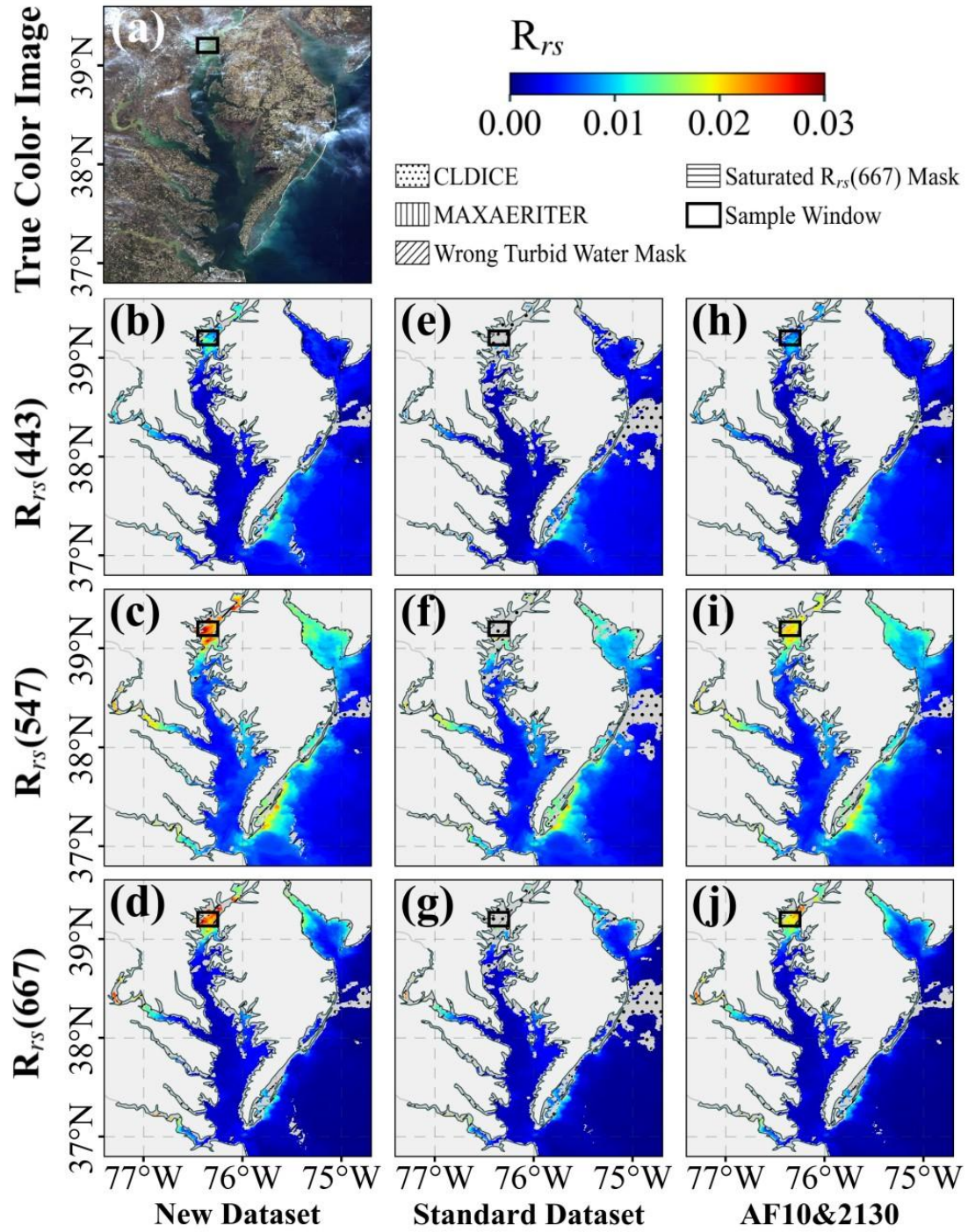


Figure 10: Comparison of R_{rs} (units: sr^{-1}) derived from single-scene MODIS Aqua observations over the Chesapeake Bay (2019-02-05). The top image shows a true color composite (panel a), followed by R_{rs} at 443, 547, and 667 nm from the new dataset (panels b–d), from the standard dataset (panels e–g) and from the AF10&2130 configuration (panels h–j). Invalid pixels are marked with hatched patterns based on quality control masks. Black rectangles mark the sample windows where average R_{rs} values are calculated and compared quantitatively in the text.

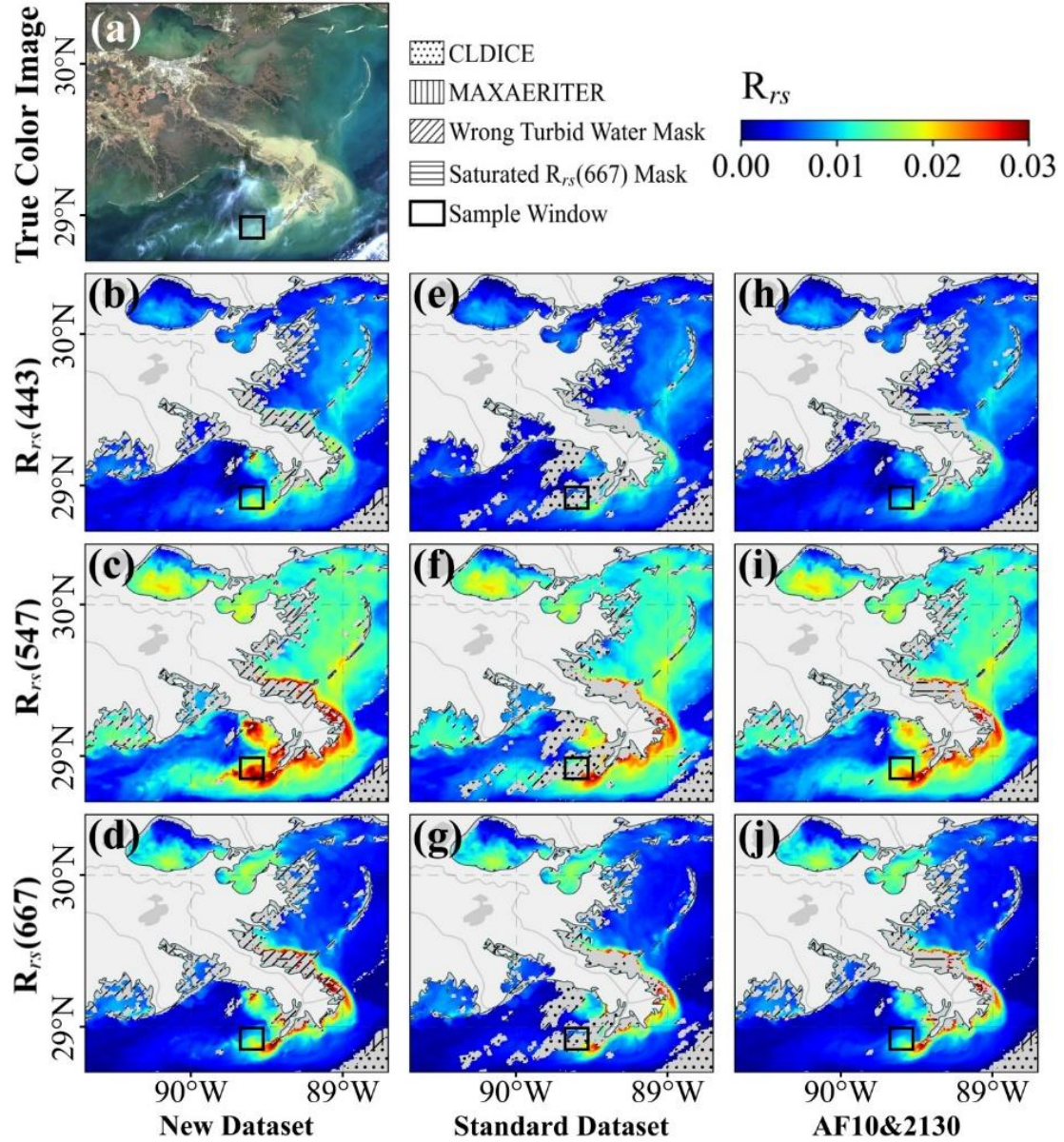


Figure 11: Comparison of R_{rs} (units: sr^{-1}) derived from single-scene MODIS Aqua observations over the Mississippi River Estuary (2021-01-13). The top image shows a true color composite (panel a), followed by R_{rs} at 443, 547, and 667 nm from the new dataset (panels b–d), from the standard dataset (panels e–g) and from the AF10&2130 configuration (panels h–j). Invalid pixels are marked with hatched patterns based on quality control masks. Black rectangles mark the sample windows where average R_{rs} values are calculated and compared quantitatively in the text.

Therefore, for the scenes analyzed, while an updated mask alone would yield a similar quantity of valid pixels, their accuracy would differ substantially. The revised section demonstrates the integrated success of our framework: it recovers data and ensures its accuracy, advancing beyond approaches that address coverage or bias in isolation. We believe these revisions have significantly strengthened the manuscript's analysis and directly address your valuable comment.

3. On a minor note. Why just the coast waters of China and the United States? Why not the coastal waters of the global ocean?

Response: Thank you for your valuable comment. The selection of the coastal waters of China and the United States was a deliberate and strategic choice, driven by our core objective: to rigorously develop and validate an improved processing framework specifically for optically complex coastal waters where standard global products consistently underperform. The coastal waters of China and the US serve as ideal, high-impact testbeds for this challenge. They encompass a wide spectrum of the most demanding conditions, from the highly turbid waters and frequent absorbing aerosols of China's coast to the diverse optical environments and anthropogenic influences along the US coast. Successfully generating a high-quality R_{rs} product in these exceptionally challenging regions provides the most robust and convincing proof-of-concept for our framework's effectiveness and practicality.

We have revised the Introduction (Section 1) to better justify this regional focus: **“The framework is applied to generate a consistent, high-quality MODIS R_{rs} dataset (2003-2022) for two strategically selected regions: the coastal waters of China and the United States (US). These regions together provide representative cases for the optical challenges in coastal zones—from the highly turbid waters and frequent absorbing aerosols of China's coast to the diverse optical environments and significant anthropogenic influences along the US coast (Charlson et al., 1992; Wang et al., 2017), and serve as a rigorous testbed to validate the framework's robustness under the most demanding conditions.”**

We agree that a global application is the ultimate goal. The significant improvements documented here establish a solid foundation for future expansion to other coastal regions and eventually a global scale. We thank the reviewer for highlighting this vital direction.

A few minor specifics:

1. Line 19: “(slope: 1.00 ± 0.08)”. Statistically, the slope alone is insufficient to indicate error or uncertainty or accuracy.

Response: Thank you for your insightful comment. We entirely agree that a slope of 1.00, while indicative of reduced systematic bias, is insufficient as a standalone measure of total accuracy or precision. In the revised manuscript, we have clarified that the slope is used specifically to demonstrate the mitigation of the pervasive systematic underestimation observed in the standard product. We have updated the Abstract to present the slope within the context of a broader suite of validation metrics, ensuring it is not framed as the sole indicator of performance. The revised text reads: **“Compared to the NASA standard MODIS Aqua R_{rs} products, the improved framework achieved a significant reduction in systematic bias. Multiple validation metrics support this, with the restoration of the mean regression slope to unity (1.00 ± 0.08 vs. 0.90 ± 0.06) across eight bands being the most prominent. Performance gains**

are particularly pronounced in the 488–555 nm range, which is most relevant to water color applications.”

2. Line 34, “Ioccg” should be “IOCCG”.

Response: Thank you for your observation. “Ioccg” has been corrected to **“IOCCG”** throughout the manuscript.

3. Line 148, “located over open ocean ... were excluded.” What is the criterion to identify such “open ocean” waters?

Response: Thank you for raising this point. To be more precise, here the excluded observations were those falling outside the specific spatial coverage of the global gridded aerosol models. To clarify this criterion, we have revised the manuscript as follows: **“To ensure compatibility with the coverage of the global gridded aerosol models and focus on coastal conditions, observations located outside the spatial domain of the global gridded aerosol models or over inland waters were excluded.”**

4. Fig. 3. It is necessary to include units.

Response: Thank you for pointing out this omission. The units for remote sensing reflectance have been added to the axes of Figure 3. The x-axis and y-axis now clearly state: **“*In situ* $R_{rs}(\text{wavelength})$ (sr^{-1})”** and **“MODIS $R_{rs}(\text{wavelength})$ (sr^{-1})”**, respectively.

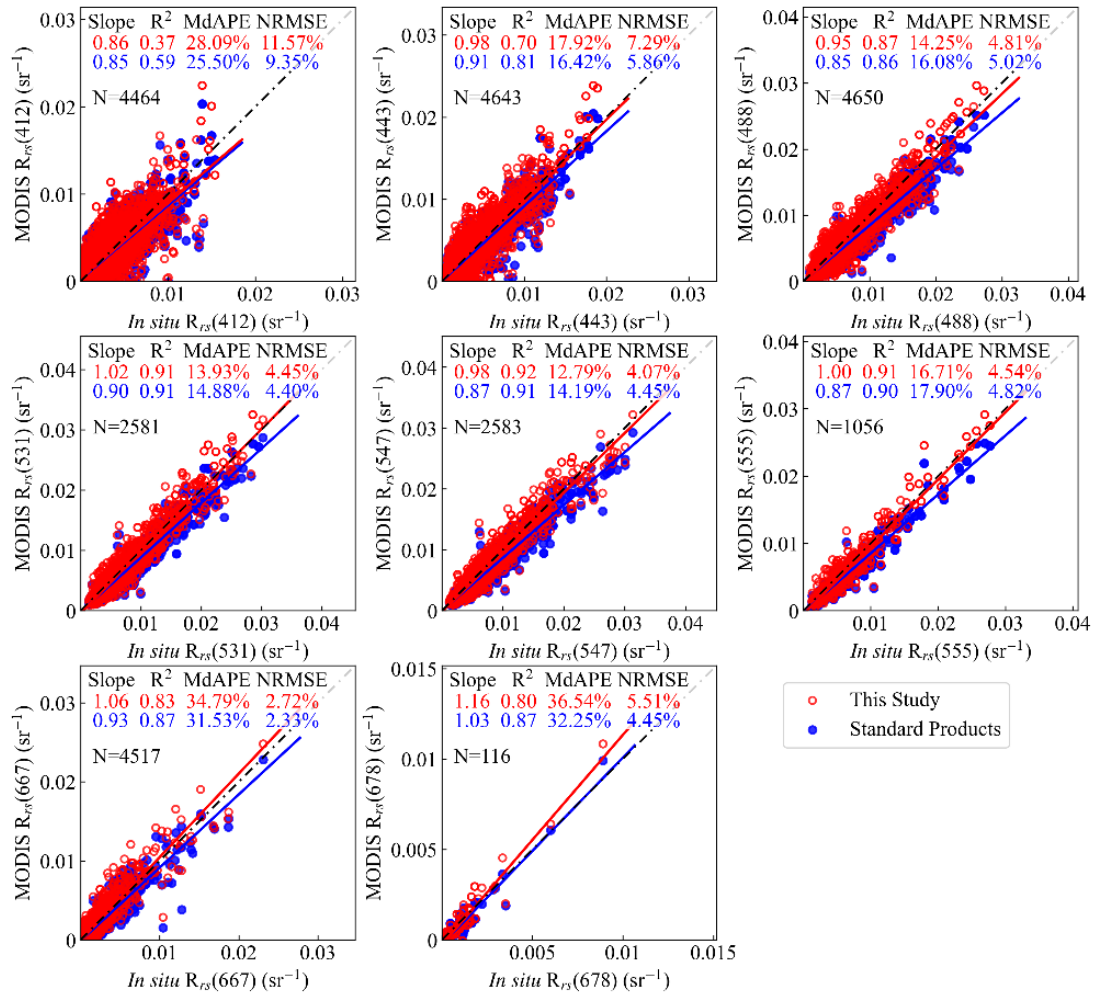


Figure 3: Scatter plots showing the validation of R_{rs} (unit: sr^{-1}) from this study and from NASA standard MODIS Aqua products against in situ measurements across eight spectral bands. Accuracy metrics and the number of match-ups (N) are annotated.

We hope these revisions and the point-by-point responses above have satisfactorily addressed your concerns. Thank you again for your insightful comments, which have significantly helped improve the clarity and quality of our manuscript.

References

- Charlson, R. J., Schwartz, S. E., Hales, J. M., Cess, R. D., Coakley, J. A., Hansen, J. E., and Hofmann, D. J.: Climate Forcing by Anthropogenic Aerosols, *Science*, 255, 423-430, 1992.
- Fan, Y., Li, W., Chen, N., Ahn, J.-H., Park, Y.-J., Kratzer, S., Schroeder, T., Ishizaka, J., Chang, R., and Stamnes, K.: OC-SMART: A machine learning based data analysis platform for satellite ocean color sensors, *Remote Sensing of Environment*, 253, 112236, <https://doi.org/10.1016/j.rse.2020.112236>, 2021.

Hu, C., Feng, L., Lee, Z., Davis, C. O., Mannino, A., McClain, C. R., and Franz, B. A.: Dynamic range and sensitivity requirements of satellite ocean color sensors: learning from the past, *Appl. Opt.*, 51, 6045-6062, 10.1364/AO.51.006045, 2012.

Liu, J., Yan, W., Chen, Z., and Lu, J.: Sediment sources and their contribution along northern coast of the South China Sea: Evidence from clay minerals of surface sediments, *Continental Shelf Research*, 47, 156-164, <https://doi.org/10.1016/j.csr.2012.07.013>, 2012.

Lv, J., Feng, L., and Zhao, D.: Validation of Global Gridded Aerosol Models in Inland/Coastal Water Atmospheric Correction for MODIS, VIIRS, and Landsat, *IEEE Transactions on Geoscience and Remote Sensing*, 62, 1-10, 10.1109/TGRS.2024.3427836, 2024.

Steinmetz, F., Deschamps, P.-Y., and Ramon, D.: Atmospheric correction in presence of sun glint: application to MERIS, *Opt. Express*, 19, 9783-9800, 10.1364/OE.19.009783, 2011.

Wang, Y., Wang, J., Levy, R. C., Xu, X., and Reid, J. S.: MODIS Retrieval of Aerosol Optical Depth over Turbid Coastal Water, *Remote Sensing*, 9, 595, 2017.

Xu, Y., Zhou, S., Hu, L., Wang, Y., and Xiao, W.: Different controls on sedimentary organic carbon in the Bohai Sea: River mouth relocation, turbidity and eutrophication, *Journal of Marine Systems*, 180, 1-8, <https://doi.org/10.1016/j.jmarsys.2017.12.004>, 2018.

Zhang, M., Ma, R., Li, J., Zhang, B., and Duan, H.: A Validation Study of an Improved SWIR Iterative Atmospheric Correction Algorithm for MODIS-Aqua Measurements in Lake Taihu, China, *IEEE Transactions on Geoscience and Remote Sensing*, 52, 4686-4695, 10.1109/TGRS.2013.2283523, 2014.

Zhao, D., Feng, L., and He, X.: Global Gridded Aerosol Models Established for Atmospheric Correction Over Inland and Nearshore *Journal of Geophysical Research: Atmospheres*, 128, e2023JD038815, <https://doi.org/10.1029/2023JD038815>, 2023.