

Author's response to the reviewers

Response to reviewer RC1

General Comments

This manuscript investigates the dynamics of water transparency across diverse marine and coastal environments, using satellite-derived Secchi disk depth (ZSD) estimates from the Copernicus Marine Service over the 1998-2022 period. The study combines seasonal and interannual analysis in transparency in relation to environmental drivers, assesses the strength and reliability of long-term trends and the role of observational uncertainty, and finally examines the influence of spatial resolution on trend characterization.

The manuscript is well structured and well written. It fits well within the scope of the Copernicus Ocean State Report. The study provides a valuable contribution by highlighting how product uncertainties and spatial resolution affect trend detection in coastal and oceanic environments.

However, the manuscript requires major revisions. Several aspects need clarification or deeper justification, particularly regarding methodology, traceability, dataset limitations, uncertainty treatment, and interpretation of results. Please, find below the review to consider:

1. Novelty and scientific contribution

The integration of observational uncertainty into the trend detection process through a Monte-Carlo approach is particularly relevant, making it a valuable methodological contribution. The study also benefits from a multi-regional design which enhances the relevance of the results. Finally, the direct comparison between 1km and 4km Copernicus ocean-colour products provides useful insights regarding the influence of spatial resolution on transparency estimates and trend detection.

However, the manuscript would benefit from presenting its scientific novelty more clearly. In particular, the authors should explicitly state what is new compared with previous assessments of water transparency, including Morel et al. (2007), the Copernicus QUID by Garnesson et al., and previous trend studies (e.g., Kwiatkowska, Werdell, IOCCG reports). It would also be useful to clarify whether

comparable trend analyses that take uncertainty into account already exist in literature, and how the proposed Monte-Carlo framework advances current practice. More generally, the introduction should more sharply frame the specific scientific contribution of this study, distinguishing it from prior work and highlighting its added value.

We thank the reviewer for this comment and the opportunity to clarify the scientific novelty of our work. Our study does not aim to develop a new method for estimating water transparency; we rely on the ZSD variable from Morel et al. (2007) as provided in the Copernicus products. What is new in our approach lies in the statistical treatment of trend detection: while previous studies have assessed temporal trends and occasionally reported uncertainties associated with variability (e.g., due to oceanographic variability or the observation sampling), they did not incorporate the per-pixel measurement error of the satellite products themselves.

By explicitly propagating these pixel-level uncertainties through a Monte-Carlo framework, our study provides a more rigorous assessment of the confidence in the detected trend slopes. This allows for a better quantification of the impact of measurement error on trend detection and represents the main added value of our work compared with prior assessments.

➔ **Corresponding lines in the revised manuscript : 32-37, 45-54**

1. Scientific approach and methods

The methodological choices constitute clear strengths. The use of the Seasonal Mann–Kendall test combined with the Theil–Sen slope estimator is well suited for non-parametric climatological trend detection. The Monte-Carlo propagation of uncertainties is also methodologically robust, providing a more realistic assessment of trend robustness than standard single-run approaches. In addition, the incorporation of river discharge data complements the analysis, and offers valuable context for interpreting seasonal and interannual variability in transparency, particularly in estuarine systems.

A first point which requires clarification concerns the treatment of uncertainty. The manuscript uses the relative uncertainties reported in the Copernicus QUID and interprets them as Gaussian noise when generating synthetic time series to reflect the inherent uncertainty of the observations. The authors should clarify whether the uncertainties are intended to represent measurement errors, random noise, systematic bias, or a mixture of multiple sources of discrepancy. The authors should also justify why a Gaussian assumption is appropriate in this context. This clarification is important, as it affects how uncertainty should be propagated and interpreted in trend detection.

We thank the reviewer for this important point and for the opportunity to clarify our treatment of uncertainty.

The relative uncertainties reported in the Copernicus QUID are intended to represent a combination of measurement errors and random discrepancies arising from the comparison of satellite products with in-situ reference datasets. These uncertainties therefore reflect the aggregated effect of multiple error sources rather than a single systematic bias.

In the Copernicus processing chain, uncertainties are propagated through successive steps including reprojection, multi-sensor merging, and temporal averaging. The assumption of Gaussian noise is motivated by the fact that these uncertainties result from the additive contribution of several independent error components introduced during the comparison with in-situ observations and subsequent processing steps. Under such conditions, the Gaussian approximation is statistically justified and commonly adopted, following the central limit theorem.

➔ **Corresponding lines in the revised manuscript : 99-101, 109-111**

A second point requiring further detail concerns the implementation of the Monte-Carlo simulations. The manuscript should specify which uncertainty metric (quality flags, per-pixel relative uncertainty, global RMS...) is used and whether these uncertainties vary seasonally or geographically. The choice of 5000 simulations should also be justified, ideally supported by a convergence analysis showing that trend statistics stabilize beyond a certain number of iterations. In addition, it is important to clarify how gaps in the time series are handled in the Monte-Carlo simulations, as inconsistent treatment of missing data could introduce artificial variability.

We thank the reviewer for these comments and provide additional details on the implementation of the Monte-Carlo simulations. The uncertainty metric used in this study is the per-pixel relative uncertainty provided directly in the Copernicus products. These uncertainty estimates are derived from comparisons with in-situ reference data and/or from the multi-sensor merging model, as documented in the Copernicus QUID. As a result, they are inherently linked to the magnitude of the ZSD variable and may exhibit moderate seasonal or geographical variability.

This variability is not expected to significantly affect our results, as uncertainty is explicitly accounted for within the trend detection framework itself. The Monte-Carlo approach propagates the reported uncertainty into the synthetic time series, thereby allowing the trend significance to be evaluated in the presence of realistic, observation-dependent noise. Consequently, potential spatial or temporal variations in uncertainty are naturally incorporated into the trend statistics rather than treated as an external or fixed perturbation.

The choice of 5000 Monte-Carlo realizations was guided by a convergence analysis, which showed that the estimated trend statistics (slope and significance) stabilize beyond this number of iterations. For the sake of conciseness and to limit the length of the manuscript, the results of this convergence test were not included, but the selected number of simulations ensures robust and stable estimates.

Finally, we agree with the reviewer that inconsistent handling of missing data could introduce artificial variability in Monte-Carlo simulations. In the present study, this issue does not arise, as the analysis is based on monthly aggregated time series with no temporal gaps.

→ **Corresponding lines in the revised manuscript : 101-107, 111-114**

The third point concerns the spatial and temporal aggregation. Monthly averaging of daily data is reasonable, but the use of the quadratic mean to aggregate uncertainties deserves explicit justification. The authors should explain why RMS was preferred over other averaging schemes and discuss the implications of this choice for variability and trend estimation.

We thank the reviewer for raising this important point regarding the spatial and temporal aggregation of uncertainties. The use of the quadratic mean (RMS) to aggregate uncertainties implicitly assumes that individual pixel uncertainties are independent. As the reviewer rightly points out, this assumption is not strictly valid in our case. The uncertainty estimates provided in the Copernicus products are partly derived from comparisons against the same in-situ reference datasets, and the observations originate from the same sensor, which introduces a degree of correlation between pixel-level uncertainties.

In principle, a rigorous aggregation of uncertainties would therefore require explicit knowledge of the covariance structure between pixels. However, such information is not available and would be difficult to estimate robustly. In this context, an alternative and conservative approach is to consider a worst-case scenario in which uncertainties are fully correlated across the study area. Under this assumption, uncertainty aggregation reduces to an arithmetic mean rather than an RMS, leading to larger aggregated uncertainty estimates.

We acknowledge that the RMS-based aggregation may underestimate the effective uncertainty if correlations are strong. We therefore consider the reviewer's remark highly relevant and are currently evaluating an updated version of the analysis using a fully correlated assumption. While this approach is expected to increase the magnitude of the aggregated uncertainties, the temporal structure and relative distribution of uncertainty along the time series are expected to remain similar. As a consequence, the main conclusions regarding trend detection are unlikely to be qualitatively altered, although the confidence levels may be more conservative.

→ **Corresponding lines in the revised manuscript : 84-87. That part also impacted some of the results and interpretations of the trend analysis, especially regarding the Gironde estuary (l. 176-188)**

The fourth point concerns the pixel selection criterion based on a 30% revisit threshold. This filter strongly affects the structure of the time series, particularly in regions with frequent cloud cover. The rationale for selecting 30% should be explained, and the authors may consider presenting a sensitivity analysis to demonstrate how results change with different thresholds (e.g., 20%, 40%, 50%). This would help assess the robustness of the trends to data availability.

We thank the reviewer for this insightful comment and constructive suggestion.

The pixel selection criterion is indeed a trade-off between temporal data availability and spatial representativeness of the study area. A 30% revisit threshold was chosen to ensure sufficient temporal coverage for reliable trend estimation while maintaining a representative spatial coverage of the North Atlantic coastal region, which is frequently affected by cloud cover. Pixels available less than 30% of the time were therefore excluded to avoid trends driven by very sparse observations.

We assessed the impact of this threshold on the estimated trends by comparing results obtained with and without the 30% filtering criterion. In both cases, the resulting trends remain non-significant, indicating that the choice of the threshold does not substantially affect the conclusions at the basin scale. While a full sensitivity analysis using multiple thresholds (e.g., 20%, 40%, 50%) was not conducted, the comparison between filtered and unfiltered datasets indicates that the main conclusions are not sensitive to the chosen threshold.

Furthermore, for the analysis of local trends, all retained pixels exhibit an availability greater than 30%, such that the revisit threshold does not influence the local trend estimates. These results suggest that the identified trends are robust to reasonable variations in data availability.

We will clarify this point in the revised version of the manuscript.

→ **Corresponding lines in the revised manuscript : 89-92**

Finally, the manuscript would benefit from a more explicit discussion of potential confounding factors influencing long-term trends in satellite-derived Secchi disk depth. Changes in transparency estimated from satellite data may result from changes in suspended particulate matter, CDOM, chlorophyll concentrations, or combinations of these components. In addition, atmospheric correction artifacts, shifts in sensor calibration, and the heterogeneity inherent in a multi-sensor time series spanning 25 years can introduce non-negligible biases. Addressing these

factors would strengthen the interpretation of the results and place the conclusions in the context of the known limitations of ocean colour climate data records.

We thank the reviewer for this important comment, which highlights key limitations in the interpretation of long-term satellite-derived water transparency trends.

The evolution of remotely sensed water transparency and Secchi disk depth is indeed influenced by multiple, potentially confounding factors. Variations in suspended particulate matter, coloured dissolved organic matter (CDOM), and chlorophyll-a concentration can all contribute to observed changes in transparency. In the revised manuscript, we will expand the discussion to more explicitly address these co-varying factors and their potential influence on the observed trends.

Regarding the consistency of the long-term time series, atmospheric correction and quality control rely on the OC5 methodology (Gohin et al., 2002), which was specifically developed for optically complex coastal waters. This approach aims to maximise spatial and temporal coverage while maintaining a robust level of data quality, as further demonstrated in Garnesson et al. (2019).

The multi-sensor time series is constructed by accounting for sensor-specific uncertainties during the merging process. To assess the potential impact of sensor changes on trend estimates, Figure 3 illustrates the contribution and lifespan of each sensor within the ZSD time series. No systematic relationship is observed between sensor transitions and the estimated trends, suggesting that sensor heterogeneity does not introduce a significant bias in the trend analysis.

→ Corresponding lines in the revised manuscript : 124-153

1. Results, discussion, and interpretation

The results and discussion sections show several strengths: seasonal variability is clearly illustrated across regions, the relationship with river discharge is convincing in the Gironde and Loire estuaries, and the comparison between 1 km and 4 km products provides useful insight into the effect of spatial resolution. However, several aspects need clearer presentation or deeper analysis.

The quality and clarity of several figures should be improved. In Figure 1, the zoomed panel for Mayotte appears blurry and should be replaced with a higher-resolution extract. Figure 3 would benefit from a more detailed legend: the four sub-plots should be clearly described, the different curves (ZSD, river discharge) explicitly identified, and the distinction between trends computed with and without uncertainties made visually clearer, as it is currently difficult to see. In Figures 3 and 4, the placement of subplot labels (a, b, c...) is confusing and should be adjusted for readability, and the y- axis label is partially cut off and needs to be

corrected. These adjustments would greatly improve the interpretability of the figures.

We thank the reviewer for these constructive comments on the figures. We agree that several aspects of the figure quality and layout can be improved. We will revise the figures to enhance their readability and clarity, including improving resolution where needed, clarifying legends, and adjusting subplot labeling and axis formatting. These changes will help make the figures easier to interpret.

→ ***The figures were updated in the revised version of the paper***

The interpretation of trends would benefit from additional support, by including spatial robustness maps, showing for each pixel the proportion of simulations in which the Seasonal Mann–Kendall test detects a significant trend. Such maps would provide a spatially explicit measure of trend reliability and would allow readers to distinguish between areas where the signal is robust and those where trends are highly sensitive to observational uncertainty. This would help clarify whether detected trends represent broad, coherent regional patterns or whether they are restricted to a few isolated pixels.

We thank the reviewer for this constructive suggestion. Mapping the consistency of trend detection across Monte-Carlo realizations at the pixel level would indeed offer additional insight into the spatial stability of the inferred trends. However, producing and interpreting these additional diagnostics would require a dedicated pixel-scale analysis and substantially extend the scope of the present study, which is focused on regional-scale temporal trends. For this reason, we believe that including this analysis goes beyond the current framework of the manuscript. We nevertheless fully agree on its relevance and will explicitly mention this approach as a perspective for future work.

The impact of spatial resolution could also be more supported. The observation that the 4 km product exhibits more significant trends while potentially smoothing fine-scale variability is interesting. Including spatial maps comparing slopes derived from the 1 km and 4 km datasets would help illustrate how resolution affects trend magnitude and spatial patterns.

In a similar way to our response to the previous comment, we agree that spatial maps comparing trend slopes from the 1 km and 4 km products would provide useful insight into the effects of resolution on trend magnitude and spatial patterns. However, performing this analysis goes beyond the scope of the current study and will be considered as a perspective for future work.

Finally, the physical interpretation of the “counter-intuitive” seasonal pattern for Mayotte, should be discussed with plausible hypotheses (e.g., lagoon sediment resuspension, cyclone-driven turbidity seasonality, tidal mixing, coral reef processes) to situate the findings within known regional processes.

The Discussion section will be revised to better contextualise the seasonal pattern observed for Mayotte by discussing plausible physical and biogeochemical processes known to affect water transparency in the region.

➔ ***Corresponding lines in the revised manuscript : 135-138***

Citation: <https://doi.org/10.5194/sp-2025-21-AC1>

Response to reviewer RC2

GENERAL COMMENTS

This paper investigates the seasonal and interannual variability of satellite-derived Secchi Disk depth (ZSD) over the period 1998–2022, comparing different environmental settings ranging from highly turbid estuarine systems to oligotrophic tropical and offshore waters. The variability is examined in relation to environmental drivers, with particular emphasis on estimates of relatively long-term trends and their associated uncertainty. This is certainly of interest for the report on the state of the oceans. On the other hand, several aspects of the methodology and data processing need to be clarified and, in some cases, better explained and justified. This is particularly true for the choice of the method adopted for the slope determination (why Theil–Sen slope estimator instead of a robust least square approach?) and the implementation of the Montecarlo approach.

We thank the reviewer for this comment and the opportunity to clarify our methodological choices. The Theil–Sen slope estimator was chosen because it is a non-parametric method that makes no assumptions about the distribution of residual errors and is less sensitive to outliers compared with ordinary least squares (OLS) regression. This property is particularly useful for trend analysis in environmental time series, where occasional extreme values can otherwise bias the slope estimate.

A known limitation of the Theil–Sen estimator is its sensitivity to missing values. In our study, this is not an issue, as we work with monthly aggregated time series that contain very few, if any, missing values.

➔ **Corresponding lines in the revised manuscript : 95-98**

The discussion about the impact of the spatial resolution should be more convincing. The “systematic offset was observed, with mean ZSD values of 21.67 m for the 1 km product and 22.50 m for the 4 km product” does not seem to be statistically significant also considering 1.40 m and 1.44 m uncertainty levels for the 1 km and 4 km products. In my view, the same holds also for the slope estimates derived from the 1 km and 4 km resolution datasets, where the difference between 0.9 cm yr⁻¹ and 2.7 cm yr⁻¹ does not seem to imply substantial consequences, despite the statistical significance indicated by the p-value tests. Probably, two maps showing the spatial distribution of the slopes at 1 km and 4 km resolution would help to qualitatively understand the consequences of spatial resolution.

We thank the reviewer for this comment and agree that the formulation in the manuscript regarding the impact of spatial resolution was not entirely appropriate. We

will revise the text to better reflect the fact that, although small differences are observed between the 1 km and 4 km products in mean ZSD values and slope estimates, these differences are likely not of substantial practical consequence given the magnitude of the uncertainties.

We also acknowledge the suggestion to include spatial maps comparing slopes at 1 km and 4 km resolution. Such maps would indeed provide a more detailed view of how spatial resolution affects both the magnitude and the spatial distribution of the detected trends. While this would be valuable, producing and interpreting these spatial comparisons requires a dedicated analysis that extends beyond the regional-scale focus of the present study. We therefore consider this as an interesting perspective for future work, which could complement the current results.

SPECIFIC COMMENTS

Point "5" in Figure 1 is somewhat misleading. I understand that it indicates the portion of the North Atlantic shown in the figure, rather than a single point. I suggest that the meanings of 1, 2, 3, 4, and 5 should also be explained in the Figure 1 caption.

We thank the reviewer for this comment. We confirm that point "5" refers to the North Atlantic region rather than a single point. We will clarify this in the Figure 1 caption and provide explanations for points 1 through 5 to improve figure readability.

→ The figures were updated in the revised version of the paper

ZSD is derived using the algorithm of Morel et al. 2007. The Morel ZSD algorithm is valid for case-1 waters. Could the authors provide a justification for applying it in coastal areas 1, 2, and 3?

We thank the reviewer for this comment regarding the applicability of the Secchi disk depth (ZSD) algorithm of Morel. We acknowledge that the coastal waters considered in this study are predominantly Case-2 waters. However, there is no strict or universally accepted separation between Case-1 and Case-2 waters, as the transition between the two regimes is generally continuous.

To characterize the dominant optical regime, we used an absorption-based criterion, relying on the ratio $(a_w + a_{ph})/a_{tot}$, with values above 50% indicating conditions closer to Case-1 waters. In this ratio a_w account for the water absorption, a_{ph} account for the absorption due to phytoplankton and a_{tot} the total absorption. This analysis confirms that most of the study areas correspond to Case-2 conditions (see the figure in the attached file).

Also, we would like to underline that there is currently no consensus on a satellite-based ZSD algorithm specifically designed for coastal waters that is suitable for analysing long-term ocean colour archives (1997–present). Although the Morel algorithm was originally

developed for open-ocean waters, it is based on fundamental bio-optical relationships and remains appropriate for monitoring relative changes in water transparency. While absolute ZSD values may differ in coastal waters, the temporal evolution is expected to remain robust when the algorithm is applied consistently over time.

→ **Corresponding lines in the revised manuscript : 78-80**

The quality of the figures is generally rather poor. In some cases, the y-axis label is not visible (for example, in Figure 3d) and legends (a), (b), (c), (d) seem to be part of the title of the next figure. It becomes clear that these are the legends from the previous figure only when reading “(e) Mayotte, 4 km” at the bottom of the last panel.

In the figures, also the graphical representation of the linear slope is quite poor. Plotting the annual average together with the linear fit may make the increasing or decreasing trend more evident (if present).

We thank the reviewer for these helpful comments. We take note of the issues regarding axis labels, panel legends, and the representation of the linear slopes. We plan to revise the figures to improve their overall clarity and readability, including clearer labeling and more effective visualization of trends alongside the annual averages.

→ **The figures were updated in the revised version of the paper**

Already at 50° N, from mid-November to mid-January, the maximum solar elevation drops below 20°, and ocean color data are not available. How is the polar night handled in this work? Figure 1, referring to January 2020 do not have data voids north of 50° N. Are these data voids interpolated using data of previous and following months?

We thank the reviewer for raising this important point regarding data availability at high latitudes during winter.

In this study, only pixels with a revisit frequency of at least 30% over the 1998–2022 time series are retained. This criterion ensures that selected pixels have a relatively high and consistent observation frequency over the full record, despite seasonal gaps in data availability.

Periods of polar night or low solar elevation, during which ocean colour observations are not available, are not interpolated using data from previous or subsequent months. These data gaps are treated as part of the natural seasonal cycle. The trend analysis is therefore remains robust, as pixels with insufficient temporal coverage or excessive uncertainty do not contribute significantly to the inferred trends.

This clarification will be added to the revised manuscript to avoid any ambiguity regarding the handling of polar night conditions.

→ **Corresponding lines in the revised manuscript : 87-92**

Page 3, line 65-66. One read “Prior to analysis, the daily satellite products were aggregated into monthly composites, using the arithmetic mean for ZSD values and the quadratic mean for the associated uncertainties.” I agree on the opportunity to aggregate the data into monthly composite but the alternate use of quadratic mean and arithmetic mean must be justified.

We thank the reviewer for this comment. The monthly aggregation of daily satellite data using the arithmetic mean for ZSD values is straightforward, as it represents the average transparency over the month. For the associated uncertainties, we used the quadratic mean (RMS) to aggregate pixel-level uncertainty estimates across the study area.

We note, however, that using the RMS implicitly assumes that individual pixel uncertainties are independent. In practice, this is not strictly the case, since the uncertainties originate from comparisons to the same in-situ datasets and from the same sensor, introducing correlations between pixels. A fully rigorous aggregation would require knowledge of the covariance structure, which is not available and would be difficult to estimate robustly.

A conservative alternative is to consider a worst-case scenario in which all pixel uncertainties are fully correlated; under this assumption, aggregation reduces to a simple arithmetic mean, which would result in slightly higher aggregated uncertainties. We plan to update our analysis using this arithmetic mean approach for uncertainty aggregation. This will likely increase the magnitude of the aggregated uncertainties, and may have some impact on the temporal distribution of uncertainty and trend estimates, although we expect the overall patterns to remain broadly similar.

→ **Corresponding lines in the revised manuscript : 84-87. That part also impacted some of the results and interpretations of the trend analysis, especially regarding the Gironde estuary (l. 176-188)**

Page 3 lines 69-69. “For the Atlantic study sites, only pixels with at least 30% revisit coverage over the entire time series were retained to ensure sufficient temporal consistency”. First, why was a threshold of 30% chosen? Second, is this 30% of the data reasonably well distributed so as to avoid bias toward the beginning or the end of the time series?

We thank the reviewer for this comment and for raising these important points.

The choice of a 30% revisit coverage threshold results from a compromise between temporal data availability and spatial representativeness of the Atlantic coastal study area, which is frequently affected by cloud cover. This threshold ensures that each retained pixel contains a sufficient number of observations to allow meaningful trend

estimation, while avoiding excessive spatial data loss that would occur with more restrictive thresholds.

Regarding the temporal distribution of valid observations, we note that the exact distribution of revisit occurrences over the time series is not a critical factor in the present analysis. The trend estimation explicitly accounts for observation uncertainty, which is directly linked to data availability. Pixels with fewer valid observations are associated with higher uncertainty, while pixels with more frequent revisits exhibit lower uncertainty. This uncertainty propagation is reflected in the confidence level of the estimated trends.

Consequently, potential uneven temporal sampling (e.g., preferential sampling at the beginning or end of the time series) does not artificially strengthen the inferred trends, but rather translates into reduced statistical confidence when data coverage is limited. This approach ensures that the robustness of the trends is evaluated in a manner consistent with the actual information content of the time series.

Accordingly, the Discussion section will be revised to moderate the interpretation of the results and to more explicitly reflect the varying levels of confidence associated with data availability.

→ *Corresponding lines in the revised manuscript : 87-92*

Citation: <https://doi.org/10.5194/sp-2025-21-AC2>