https://doi.org/10.5194/essd-2025-484
Preprint. Discussion started: 1 September 2025

Open Access Earth System Science Data Discussions

A centennial-scale data of suspended particulate matter over the global ocean

Jun Chen[1,2*], Cheng Fan[3*], Qianni Deng[1], Liguo Liu[3], Qianguo Xing[4], Shilin Tang[5], Wenting Quan[1], Xiaofan Wang[1]

[1]Key Laboratory of Physical Electronics and Devices, Ministry of Education, Faculty of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China
[2]State Key Laboratory of Satellite Ocean Environment Dynamics, Second Institute of Oceanography, Ministry of Natural Resources, Hangzhou, 310012, China
[3]College of Electronic Engineering, Naval University of Engineering, Wuhan, 430033, China
[4]Yantai Institute of Coastal Zone Research, Chinese Academy of Sciences, Yantai 264003, China
[5]South China Sea Institute of Oceanology, Chinese Academy of Sciences, Guangzhou, 510301, China

Correspondence to: Jun Chen: chenjun@xjtu.edu.cn and Cheng Fan: 1807111057@nue.edu.cn

Abstract: Suspended Particulate Matter (SPM) is a crucial indicator of marine nutrient transport, the carbon cycle, and land-sourced pollution. However, SPM research has been constrained by limited data. While marine remote sensing has helped fill this gap since late 20th century, pre-2000 data remains scarce, particularly for global datasets with high spatiotemporal resolution. To address the shortage of in-situ SPM observations, we developed two empirical models for estimating SPM concentrations: one based on remote sensing reflectance (R_{rs}) and the other on Secchi depth (Z_{sd}). Model evaluation shows that retrieval uncertainties do not exceed 35%. To mitigate errors in highly turbid waters (>100 mg l^{-1}), we applied data quality constraints by excluding samples above this threshold. The resulting dataset merges in-situ SPM observations, providing global coverage of SPM concentrations from 1890 to 2020. Temporal and spatial analysis reveals an overall increase in global SPM concentrations, with a sharp rise between 1920 and 1942, followed by a rapid return to pre-1920 levels. The underlying causes of this fluctuation warrant further investigation. This unique dataset holds considerable potential for marine resource management, climate modeling, human activity assessment, and environmental protection.

Keyword: global dataset; remote sensing reflectance; secchi depth; high spatiotemporal resolution

1 Introduction

Suspended particulate matter (SPM) is a key factor influencing the ecological condition of coastal environments, the efficiency of marine primary production, global biogeochemical cycles, and climate-related ocean processes (Loisel et al., 2014; Scown et al., 2023). High levels of SPM can lead to significant alterations in shoreline dynamics, reduce navigability in estuarine and harbor regions, and pose challenges for coastal sediment regulation and planning (Green et al., 2000). Moreover, by controlling the attenuation of light within the water column, SPM plays a fundamental role in shaping the underwater light regime, which in turn affects the vertical distribution of photosynthetically active radiation and the growth

https://doi.org/10.5194/essd-2025-484
Preprint. Discussion started: 1 September 2025

dynamics of phytoplankton populations (D'Sa et al., 2007). Beyond optical effects, SPM acts as a medium for transporting key biogeochemical components—including nutrients, heavy metals, and organic pollutants—and contributes to carbon cycling between oceanic and atmospheric systems (Yoshida et al., 2024). Recent investigations have emphasized that fluctuations in SPM concentrations are closely associated with both natural variability and anthropogenic climate stressors, such as intensifying storm events, rising sea levels, and enhanced glacial melt (Overeem et al., 2017; Szeligowska et al., 2024).

Understanding the influence of human activities and climate change on global sediment cycles requires long-term, consistent observations. Ship-based measurements of suspended particulate matter (SPM_{mea}), initiated in the 1930s using isokinetic and later bottle samplers, remain spatially and temporally sparse due to logistical constraints (Milliman and Meade, 1983). The advent of satellite remote sensing has since enabled broad-scale assessments of marine SPM, with numerous studies reporting widespread declines in sediment delivery from rivers, primarily linked to dam construction (Dethier et al., 2022). Yet, despite these advances, the temporal coverage of satellite observations remains limited (1984–present), leaving the evolution of sediment dynamics prior to major dam-building episodes largely unquantified (Rodriguez et al., 2020). This temporal discontinuity creates a critical barrier to resolving century-scale changes in global sediment flux. Bridging this gap necessitates the development of integrated, multi-source datasets that leverage both historical and contemporary observations, in combination with modeling approaches, to reconstruct long-term SPM variability with improved accuracy, spatial representativeness, and temporal continuity.

Water transparency, defined as the depth to which light penetrates into the water column, is typically measured using a Secchi disk—the depth at which the disk disappears from view as it is lowered into the water (Tyler, 1968). As one of the oldest and most straightforward bio-optical measurements, Secchi disk transparency has been routinely collected in oceanographic surveys since 1865 (Cialdi and Secchi, 1865) , predating systematic SPM measurements by several decades. Historically, transparency has explained a significant portion of the variability in marine bio-optical properties and effectively captured large-scale changes in ocean optical environments (Pitarch, 2020), making it a valuable indicator of both climatic and ecological dynamics (Boyce et al., 2010). Importantly, standardized Secchi disk observations can be empirically related to SPM concentrations (SPM_{zsd}) through some well-established bio-optical models (Swift et al., 2006), as transparency is closely linked to SPM dynamics and their influence on light attenuation (Preisendorfer, 1986). Given the spatial extent, temporal coverage, and historical continuity of these measurements, transparency records provide a unique opportunity to extend reconstructions of marine SPM more than a century beyond the advent of direct SPM observations, thereby bridging a critical gap in our understanding of long-term sediment dynamics.

Remote sensing reflectance (R_{rs}) serves as a fundamental parameter in ocean color remote sensing, as it underpins the derivation of a wide range of downstream ocean color products, including suspended particulate matter (Chen et al., 2015; Mao et al., 2012), as well as other biogeochemical indicators (IOCCG, 2006). Since the launch of the Coastal Zone Color Scanner (CZCS) in late 1978 (Gordon et al., 1980b), a substantial volume of R_{rs} observations has been accumulated globally. Notably, extensive field-based R_{rs} measurements are archived and made publicly accessible through repositories such as the

https://doi.org/10.5194/essd-2025-484
Preprint. Discussion started: 1 September 2025

SeaWiFS Bio-optical Archive and Storage System, SeaBASS (Werdell and Bailey, 2002). In contrast to SPM_{mea}, which requires labor-intensive and often logistically complex in-situ sampling, R_{rs} data can be acquired more efficiently and at significantly lower cost. Consequently, the availability of R_{rs} datasets far exceeds that of direct SPM measurements. Developing reliable algorithms to retrieve SPM concentrations from R_{rs} (SPM_{rrs}) would therefore be highly advantageous, enabling retrospective analysis of sediment dynamics and enhancing the spatial and temporal resolution of SPM monitoring beyond what is feasible through conventional field sampling alone.

In the era of big data, the effective integration and utilization of multisource SPM information are crucial for addressing persistent knowledge gaps concerning the response of global sedimentary processes to climate variability and anthropogenic influences in the Anthropocene. This study proposes a globally consistent SPM dataset constructed by harmonizing observations from multiple independent sources, aiming to reconstruct long-term records of oceanic SPM with enhanced temporal and spatial continuity (Fig. 1). Two empirical conversion models are introduced to derive SPM concentrations from in situ remote sensing reflectance and water transparency measurements, respectively, and the consistency and reliability of these multisource estimates are systematically assessed. The resulting dataset spans from 1890 to 2020, offering annual temporal resolution and a spatial resolution of 5°, thereby enabling the examination of decadal-scale trends in sediment variability. By bridging satellite-era and historical records, this dataset provides a robust basis for retrospective studies of sediment processes and supports improved projections of future SPM dynamics under changing environmental conditions.

2 Data, method, and techniques

2.1 SPM_{mea} data originated from bottle field measurements

Bottle-based field measurements for SPM_{mea}—particularly those using filtration and gravimetric analysis—are rooted in classical limnological and oceanographic practices that trace back to the mid-19th century (Strickland and Parsons, 1972). This method involves collecting water samples from various depths using specialized water sampling devices, such as Niskin bottles. The collected samples are subsequently filtered through pre-weighed glass fiber filters to retain suspended particulates. After filtration, the filters are dried and reweighed to determine the mass of the suspended matter.

Even though bottle-based field measurements were developed relatively early, the limited accessibility of ocean environments has resulted in most shared SPM datasets being concentrated in inland lakes and rivers, with comparatively fewer records from marine regions. In this study, SPM_{mea} data are majourly obtained from platforms such as Sea-viewing Wide Field-of-view Sensor Bio-optical Archive and Storage System (SeaBASS) and United States Geological Survey (USGS) Water Quality Portal (See Table 1). In addition, Lavigne et al. (2022) have released a set of publicly available SPM datasets, which we have also incorporated into our data collection.

https://doi.org/10.5194/essd-2025-484
Preprint. Discussion started: 1 September 2025

2.2 Empirical conceptual model for generating SPM_{zsd} data

Secchi disk depth (Z_{sd}), one of the most established bio-optical proxies, has been a cornerstone of marine monitoring since its introduction in 1865 (Cialdi and Secchi, 1865)—a period that coincides with the early development of bottle-based field measurements (Strickland and Parsons, 1972). However, the relative simplicity of transparency measurements compared to the more labor-intensive procedures required for SPM_{mea} quantification has contributed to the broader availability and greater abundance of transparency datasets.

Tyler (1968) and Preisendorfer (1986) identified Z_{sd} as an optical measure of water clarity, representing the depth at which a submerged target is no longer visible. The value of Z_{sd} is primarily governed by the composition and concentration of various light-attenuating substances (Effler et al., 2017). Numerous principal component analyses have demonstrated that Z_{sd} is inversely correlated with SPM concentration (Golubkov and Golubkov, 2024), as elevated turbidity levels reduce water transparency to the point where the Secchi disk becomes undetectable to the human eye. Given these inherent covarying mechanism, it is reasonable to hypothesize that SPM concentrations can be inferred from Z_{sd} (SPM_{zsd}) through a empirical conceptual model (termed as SPM_{zsd} model, h), such as:

$$SPM_{\mathrm{zsd}} = h(Z_{\mathrm{sd}}) \,, \tag{1}$$

Secchi disk transparency data have been collected globally for over 130 years (Boyce et al., 2010). If the conceptual model in Equation (1) is valid, it would enable a substantial improvement in the spatial and temporal resolution of historical and contemporary global SPM datasets. This, in turn, could provide crucial observational evidence for evaluating the impacts of the First and Second Industrial Revolutions on the global suspended sediment cycle, thereby contributing to the reduction of existing knowledge gaps in this field.

2.3 Empirical conceptual model for generating SPM_{rrs} data

As SPM accounts for a significant portion of the variability in water-leaving radiance (Morel and Prieur, 1977), it has been widely acknowledged as a reliable proxy for estimating SPM concentration from remote sensing reflectance (R_{rs}) across global oceanic and coastal waters. To date, three primary types of models have been developed for this purpose: empirical, semi-analytical, and analytical approaches (Shen and Verhoef, 2010).

Commonly, semi-analytical and analytical models are theoretically more robust and physically interpretable due to their explicit incorporation of radiative transfer principles (Chen et al., 2014; Sun et al., 2013), but extensive practical applications have shown that empirical models frequently outperform them in terms of solution stability, predictive accuracy, and consistency with in situ observations (Allison et al., 2010; Dierssen, 2010; Sokoletsky et al., 2016). This discrepancy arises partly because the highly heterogeneous nature of suspended particulate matter—characterized by its diverse composition, size distribution, and scattering behaviors (Wozniak et al., 2010)—poses significant challenges for accurate parameterization within semi-analytical and analyticsal frameworks. In contrast, such complexity is inherently embedded in the observed R_{rs}

https://doi.org/10.5194/essd-2025-484
Preprint. Discussion started: 1 September 2025

signal, and when effectively exploited, empirical models are capable of capturing and leveraging this information to produce reliable estimations.

Overall, considering both practical performance and implementation efficiency, we adopt an empirical approach for estimating SPM from R_{rs}, the empirical conceptual model (termed as SPM_{rrs} model, g) of which is outlined as follows:

$$SPM_{rrs} = g[R_{rs}(412), R_{rs}(443), R_{rs}(490), R_{rs}(555), R_{rs}(670)] , \quad (2)$$

To advance the field of ocean color remote sensing, a substantial archive of R_{rs} data has been accumulated since 1991, supported by well-established data sharing frameworks. One prominent example is the SeaBASS, which remains the most widely utilized platform for the dissemination and application of R_{rs} datasets (Hooker et al., 1994). Additionally, Valente et al. (2022) have made publicly available a comprehensive dataset comprising extensive in situ R_{rs} measurements. If SPM concentrations can be accurately retrieved from R_{rs} using empirical models, these datasets could serve as a valuable resource for SPM estimation and help bridge spatiotemporal gaps in bottle-based field measurements.

2.4 Numerical simulated data

Optical properties achieved from in situ measurements are seldom free from error. Specifically, even when the most stringent protocols for remote sensing reflectance measurements are followed (Mueller et al., 2003), uncertainties in in situ radiometric measurements typically range from 3% to 5% (Hooker and Maritorena, 2000). Furthermore, Chen et al. (2025a) demonstrated that individual variability in the contrast threshold of the human eye can result in uncertainties exceeding 10% in Z_{sd} measurements. Consequently, it is essential to conduct sensitivity analyses of both the SPM_{rrs} and SPM_{zsd} retrieval models to establish robust quality control standards for the acquisition of high-precision, centennial-scale SPM datasets.

To assess the sensitivity of the SPM_{rrs} and SPM_{zsd} retrieval models to measurement noise, we designed two numerical simulation scenarios:

(1) A dataset of 6,000 transparency data points was randomly generated, uniformly distributed between 0.1 and 50 meters. Each data point was subjected to random perturbation with a uniform error not exceeding ±20%, resulting in a corresponding noise level of approximately 10% in the transparency measurements which was comparable with the disturbance associated with variations in human eye contrast (Chen et al., 2025a). The sensitivity of the SPM_{zsd} retrieval model was assessed by comparing the consistency of SPM_{zsd} derived from an error-free Z_{sd} dataset with SPM_{zsd} derived from the error-contaminated dataset.

(2) A set of 6,000 R_{rs} data points was randomly generated, with SPM values uniformly distributed between 0.1 and 500 mg/l. Each data point was perturbed by a uniform error not exceeding ±10% following approch proposed by Chen et al. (2016), resulting in an average measurement error of approximately 5% in the remote sensing reflectance values. The sensitivity of the SPM_{rrs} retrieval model was evaluated by comparing the consistency of SPM_{rrs} derived from an error-free R_{rs} dataset with SPM_{rrs} derived from the error-contaminated R_{rs} dataset.

https://doi.org/10.5194/essd-2025-484
Preprint. Discussion started: 1 September 2025

Open Access Earth System Science Data Discussions

2.5 Statistical evaluation

We used two statistics to evaluate how well the model in generation of SPM from remote sensing reflectance and transparency measurements. These statistics are the mean absolute percent difference (MAPD) and root-mean-square error (RMSE). The statistics are described as follows:

$$MAPD = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{x_{m,i} - x_{r,i}}{x_{r,i}}\right| \times 100\% \quad , \tag{3}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x_{m,i} - x_{r,i}\right)^2} \quad , \tag{4}$$

where, n is the number of elements; $x_{m,i}$ is the model-derived SPM of the i^{th} element from the R_{rs} data, and $x_{r,i}$ is the equivalent known SPM of the i^{th} element.

3 Results and discussion

3.1 Data availability from multiple sources over the global oceans

From a temporal perspective, the technological levels represented by the remote sensing reflectance, suspended particulate matter (SPM), and transparency measurement devices exhibit significant variation, leading to notable differences in the earliest available public data (Figs. 2a-2c). For instance, the earliest recorded remote sensing reflectance data dates back to 1991, while public SPM data can be traced as far back as 1966. Transparency data, on the other hand, is available from as early as 101 years and 76 years before the aforementioned two data types, respectively. These findings further confirm the fact that transparency remains, without question, the oldest optical parameter (Tyler, 1968).

In terms of data quantity, the transparency dataset comprises 583,613 samples, which is 2.4 times larger than the remote sensing reflectance dataset and 7.4 times larger than the SPM dataset (Figs. 2d-2f). Spatially, most of the SPM measurement data samples are concentrated in coastal areas, whereas the R_{rs} data samples are distributed across both coastal and open ocean regions, although their density is much lower compared to the transparency dataset. A comparison also reveals that the data samples from the Northern Hemisphere outnumber those from the Southern Hemisphere. This discrepancy is likely attributed to the fact that human activity is predominantly concentrated in the Northern Hemisphere, which houses approximately 90% of the global population (Nations, 2022).

Regarding the earliest transparency measurement dates, data from the Eastern Hemisphere is generally older than that from the Western Hemisphere (Fig. 2f). Notably, in the seas around Western Eurasia and Africa, a significant portion of the transparency measurement data dates back to before 1930, whereas most of the data from the Americas and Australia has earliest recorded measurements after 1950. Moreover, influenced by ocean accessibility and maritime technological

https://doi.org/10.5194/essd-2025-484
Preprint. Discussion started: 1 September 2025

advancements, the earliest measurement dates gradually converge from coastal areas to the open ocean. These temporal differences are crucial for the proper utilization of this dataset in relevant research.

3.2 Model initialization and evaluation for SPM retrievals

Two datasets were employed for model initialization and evaluation: one dataset contained synchronized measurements of Z_{sd} and SPM, while the other consisted of synchronized measurements of R_{rs} spectra and SPM. Both datasets were divided into two subsets: approximately 70% of the data was used for model calibration, and the remaining 30% was reserved for model validation. These datasets encompassed both clear open ocean waters and turbid coastal waters. Importantly, the dataset for the SPM_{rrs} model was notably more turbid than that for the SPM_{zsd} model. This discrepancy arises from the fact that, in extremely turbid waters, transparency values are often very low (Table 2). Even slight disturbances, such as wind or waves, can lead to substantial errors in transparency measurements. Consequently, due to the lower quality of transparency data in highly turbid waters, such samples are typically less abundant.

Fig. 3 illustrates the performance of the empirical SPM_{zsd} model in estimating SPM from transparency measurements. It was observed that the Z_{sd} value was negatively correlated with the SPM concentration, with a determination coefficient of 0.86 (Fig. 3a). This indicates that Z_{sd} could explain 86% of the variability in SPM across the global oceans, where SPM concentrations ranged from 0.5 mg l^{-1} to 120.48 mg l^{-1}. The evaluation was conducted by comparing the SPM values predicted by the SPM_{zsd} model with analytically measured SPM values from independent datasets collected globally. Based on 148 field samples, the SPM_{zsd} model was found to yield an uncertainty of 34.91% in estimating SPM from Z_{sd} measurements, with a determination coefficient of 0.71 and a slope of 1.04 (Fig. 3b). Notably, this performance was achieved without the need for any re-optimization of the model parameters.

To identify the optimal empirical SPM_{rrs} model for converting R_{rs} to SPM, we employed the non-linear iterative method proposed by Chen and Quan (2013). This method enabled us to derive the best empirical relationship between SPM and the R_{rs} spectrum. The results showed that by using the expression $X=[R_{rs}(670)+0.368R_{rs}(443)-0.463R_{rs}(412)]/[R_{rs}(490)+0.515R_{rs}(555)]$ as inputs, SPM could be accurately and linearly derived from optical data, achieving a determination coefficient of 0.87 for optical data with SPM concentrations ranging from 0.1 mg l-1 to 2239.9 mg l^{-1} (Fig. 4a). When this model was applied to an independent optical dataset with SPM concentrations ranging from 0.13 mg l^{-1} to 577.2 mg l^{-1}, the SPM_{rrs} model produced estimates with a mean absolute percentage deviation (MAPD) of 33.68%, a determination coefficient of 0.82, and a slope of 0.89 (Fig. 4b). This suggests that the SPM-R_{rs} model could account for 82% of the variability in SPM across global R_{rs} data.

Looking at the historical context of SPM estimation models, it is evident that the estimation accuracy generally falls within the range of 30% to 50% (Balasubramanian et al., 2020; Chen et al., 2015; Jiang et al., 2021; Nechad et al., 2010; Teng et al., 2025). Compared to typical SPM retrieval accuracy, both the SPM_{zsd} and SPM_{rrs} models are stable and accurate for determining SPM from optical data. These models hold considerable potential for bridging spatial and temporal gaps in long-term global SPM datasets, especially in regions where data coverage is sparse.

https://doi.org/10.5194/essd-2025-484
Preprint. Discussion started: 1 September 2025

3.3 Consistency evaluation and sensitivity analysis

The consistency between heterogeneous datasets is a crucial metric for assessing data quality in multisource data fusion. Given the inherent discrepancies between the SPM_{zsd} and SPM_{rrs} models when applied to water bodies, it is vital to evaluate the consistency of the SPM values derived from both models in order to develop robust data quality control criteria. These criteria will assist in identifying and excluding data with poor consistency from multisource data, thereby enhancing the overall dataset quality. To this end, we selected synchronized samples of Z_{sd} and R_{rs} from our field measurements and processed them through both the SPM_{zsd} and SPM_{rrs} models to derive corresponding SPM_{zsd} and SPM_{rrs} values.

As shown in Fig. 5, the overall deviation between the SPM values derived from the SPM_{zsd} and SPM_{rrs} models is 36.99%, while the corresponding determination coefficient is 0.62 (Fig. 5a). This suggests that both models are capable of producing consistent SPM estimates from Z_{sd} and R_{rs} data, but the SPM_{zsd} model tends to yield slightly higher SPM values than the SPM_{rrs} model (Fig. 5b). Particularly, the linear relationship between the two models has a slope of 1.15 and a bias of 1.639. This deviation is most prominent in samples where SPM concentrations exceed 100 mg/l, and where transparency values fall below 0.28 m, as indicated by the relationship shown in Fig. 3a.

In practice, accurately measuring transparency in highly turbid waters presents significant challenges due to a variety of factors, including wave disturbances, inconsistent operator techniques, vessel platform instability, and the environmental adaptation of human eye contrast sensitivity (Chen et al., 2025a; Pitarch, 2020). These factors contribute to significant measurement errors, especially in regions with low transparency. The SPM_{zsd} model expresses an inverse relationship between the SPM value and Z_{sd} (Fig. 3a), meaning that in low-Z_{sd} regions, where small values are divided by large ones, even minor disturbances in Z_{sd} can lead to considerable estimation errors in the SPM values. Furthermore, Doxaran et al. (2002; 2009) have demonstrated that in clear to moderately turbid waters (SPM <100 mg l^{-1}), the red and green bands of remote sensing reflectance exhibit a strong correlation with SPM concentrations. However, the successful application of near-infrared bands is contingent upon the fact that in highly turbid waters, sufficient scattering occurs to overcome the strong absorption of these bands by water (SPM > 100 mg l^{-1}), where the green and red bands become less sensitive to increases in SPM due to signal saturation (He et al., 2013). This phenomenon underscores the potential utility of the near-infrared band for remote sensing of SPM in highly turbid waters. Nevertheless, publicly available observational datasets often contain limited near-infrared data, which could lead to a significant reduction in available bio-optical data, as demonstrated in the dataset from Valente et al. (2022), if the near-infrared band is incorporated into the model inversion.

Fig. 6 illustrates the response of the SPM_{zsd} and SPM_{rrs} models to noise in field-measured R_{rs} and Z_{sd} data. Both models show strong noise tolerance ability in clear to moderately turbid waters (<100 mg l^{-1}), even though there are substantial uncertainties contained in field measurements. For example, when R_{rs} measurements contain 5% uncertainty and Z_{sd} measurements contain 10% uncertainty, the SPM_{zsd} and SPM_{rrs} models exhibit overall inconsistencies of 10.02% and 4.7%, respectively, compared to SPM values derived from error-free R_{rs} and Z_{sd} data. However, in highly turbid and extremely turbid waters (>100 mg l^{-1}), scatterplots diverging from the 1:1 line become apparent (Figs. 6a and 6c), with the model

https://doi.org/10.5194/essd-2025-484
Preprint. Discussion started: 1 September 2025

inconsistency, represented by MAPD values, increasing as SPM concentrations rise (Figs. 6b and 6d). This deviation is primarily attributed to the heightened sensitivity of the SPM_{zsd} model to noise in low-Z_{sd} regions, as well as the complex optical properties of highly turbid waters that complicate the remote sensing of SPM concentrations (D'Sa et al., 2007; He et al., 2013). Notably, in clear and moderately turbid waters, both the blue-green remote sensing reflectance and transparency values are relatively high (Chen et al., 2016), meaning that minor environmental disturbances have minimal impact, resulting in more reliable data quality.

Given the performance of both models in turbid waters and the quality of the centennial-scale SPM dataset, we recommend excluding data with SPM_{mea}, SPM_{rrs}, and SPM_{zsd} concentrations exceeding 100 mg l^{-1} from the data fusion process, as illustrated in Fig. 1.

3.4 Global centennial-scale SPM dataset

By excluding samples with SPM concentrations exceeding 100 mg l^{-1}, Fig. 7 presents the global centennial-scale SPM dataset, derived by merging SPM_{mea}, SPM_{rrs}, and SPM_{zsd} data. Among these, the Z_{sd} dataset stands out with the largest sample size and the longest temporal coverage (Fig. 1). A comparison between the Z_{sd} and the merged multisource dataset (Figs. 1c and 7a) reveals that the final dataset contains 899,255 samples—1.54 times the number of Z_{sd} samples. This increase is largely due to the contributions of SPM_{mea} and SPM_{rrs} (Fig. 7c), with the samples predominantly concentrated in coastal and some oceanic areas (Figs. 1a and 1b). It is important to note that, as R_{rs} and SPM measurement technologies were developed and became widespread much later than Z_{sd} (Cialdi and Secchi, 1865; Gordon et al., 1980a; Strickland and Parsons, 1972), the additional SPM_{mea} and SPM_{rrs} data primarily cover the past 50 years, with limited representation of earlier periods. This creates some temporal gaps in the dataset, highlighting the need for further efforts to supplement the early data.

Building on this, we analyzed the temporal variations in global SPM concentrations over nearly 130 years (Fig. 7d). The data indicate an oscillating growth pattern in SPM concentrations. Between 1890 and 1920, SPM concentrations increased slowly, but from 1920 onwards, the growth rate sharply accelerated, peaking in 1926. This was followed by a significant decline, with concentrations returning to pre-1920 levels by 1936. The reasons behind the sharp spike in SPM concentrations during the 1920-1936 period remain unclear and require further investigation. After 1942, a gradual increase in SPM concentrations continued until 2010, when the trend reversed and began to decline, persisting until 2020. The growth observed between 1942 and 2010 is likely linked to large-scale global water infrastructure projects, which led to increased coastal erosion and, consequently, higher SPM concentrations (Dethier et al., 2022; Hou et al., 2024). However, the sudden decline in concentrations after 2010, lasting over a decade, is still not fully understood and demands additional research.

While every effort has been made to collect available open datasets, this dataset represents the most extensive and temporally comprehensive global SPM dataset available in the public domain. However, it does not encompass all data accumulated globally, as many datasets are still held by individuals or organizations. Understanding the global SPM source-sink processes and their interactions with climate change and human activities requires long-term, high-resolution temporal

https://doi.org/10.5194/essd-2025-484
Preprint. Discussion started: 1 September 2025

data. The spatial distribution of the data samples is crucial for accurately reconstructing temporal trends. We believe this dataset is vital for understanding the mechanisms driving SPM movement and is currently the most comprehensive in terms of both spatial and temporal coverage. However, there is room for improvement, especially in the resolution of both spatial and temporal dimensions. We therefore urge all relevant stakeholders worldwide to make their datasets as accessible as possible to further advance this field.

In addition to optical parameters such as R_{rs} and Z_{sd}, which are closely related to SPM concentrations, other factors—such as water turbidity and sedimentation rate—also strongly correlate with SPM concentrations (Kim et al., 2025; Petus et al., 2010; Strakhovenko et al., 2023). These parameters could serve as valuable supplementary data sources for future research. We encourage scholars in this field to further explore these aspects and expand the scope of study.

4 Data availability

The synthetic optical database described in this study is publicly available at the website of https://gr.xjtu.edu.cn/en/web/chenjun/data_sharing or https://doi.org/10.5281/zenodo.16991206(Chen et al., 2025b).

5 Conclusion

SPM plays a critical role in the transport and cycling of marine chemical elements and is also a vital source of nutrients for marine phytoplankton to carry out primary production. To enhance our understanding of global climate change and the impact of human activities on the marine environment, we constructed a centennial-scale dataset of SPM across the global ocean. This dataset is of considerable scientific value, not only for advancing our knowledge of the oceanic carbon cycle but also for improving our understanding of changes in marine primary production capacity and carbon cycling efficiency. Therefore, this dataset holds significant promise for advancing related scientific fields.

Our results demonstrate that the generated centennial-scale dataset exhibits high quality. For instance, SPM data derived from Z_{sd} and R_{rs} estimations show an average error of less than 35% when compared with in situ SPM observations, making it robust enough to support large-scale global trend analyses. However, it is important to note that when SPM concentrations exceed 100 mg l^{-1}, the quality of the SPM data estimated from Z_{sd} and R_{rs} decreases significantly. For example, the consistency between the SPM_{zsd} and SPM_{rrs} models is poorer in high-concentration regions. To ensure the dataset's quality, we excluded samples where the SPM concentration exceeded 100 mg l^{-1}. Consequently, this dataset does not include water bodies with SPM concentrations exceeding this threshold. To incorporate high-concentration regions, supplementary data or methods would be required.

Overall, the centennial-scale SPM dataset marks a significant advancement in our understanding of the centennial-scale changes in SPM across the global oceans. The accuracy and stability of the dataset suggest that it is a valuable resource for understanding and predicting the dynamics of SPM concentrations on a global scale. This dataset will further contribute to

https://doi.org/10.5194/essd-2025-484
Preprint. Discussion started: 1 September 2025

improved predictions of how SPM concentrations will change under climate change. It is anticipated that the findings from this study will be of great interest to a broad spectrum of researchers, policymakers, and managers involved in the monitoring and management of aquatic ecosystems.

Acknowledgments

This study was supported by The National Key R&D Program of China (2022YFB3903003-2, Chen).

Author contributions

Conceptualization: JC. Methodology: CF and LL. Data curation: QD and QX. Writing (initial): JC. Writing (review and editing): CF, QD, LL, QX, ST, WQ and XW. Supervision: JC.

Competing interests

The contact author has declared that none of the authors has any competing interests.

Disclaimer

Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.

Financial support

This research has been supported by The National Key R&D Program of China (2022YFB3903003-2, Chen).

https://doi.org/10.5194/essd-2025-484
Preprint. Discussion started: 1 September 2025

References

Allison, D. B., Stramski, D., and Mitchell, B. G.: Empirical ocean color algorithms for estimating particulate organic carbon in the Southern Ocean, J. Geophys. Res. Oceans, 115, -, https://doi.org/10.1029/2009JC006040, 2010.

Balasubramanian, S. V., Pahlevan, N., Smith, B., Binding, C., Schalles, J., Loisel, H., Gurlin, D., Greb, S., Alikas, K., Randla, M., Bunkei, M., Moses, W., Nguyễn, H., Lehmann, M. K., O'Donnell, D., Ondrusek, M., Han, T.-H., Fichot, C. G., Moore, T., and Boss, E.: Robust algorithm for estimating total suspended solids (TSS) in inland and nearshore coastal waters, Remote Sens. Environ., 246, 111768, https://doi.org/10.1016/j.rse.2020.111768, 2020.

Boyce, D. G., Lewis, M. R., and Worm, B.: Global phytoplankton decline over the past century, Nat., 466, 591-596, 2010.

Chen, J., Cui, T. W., Qiu, Z. F., and Lin, C. S.: A three-band semi-analytical model for deriving total suspended sediment concentration from HJ-1A/CCD data in turbid coastal waters, ISPRS J. Photogramm. Remote Sens., 93, 1-13., https://doi.org/10.1016/j.isprsjprs.2014.02.011, 2014.

Chen, J., Cui, T. W., Song, Q. J., and Quan, W. T.: Estimation of total suspended matter concentration from MODIS data using a neural network model in the East China Sea, Estuar. Coast. Shelf Sci., 155, 104-113., https://doi.org/10.1016/j.ecss.2015.01.018, 2015.

Chen, J., Han, Q., Chen, Y., and Li, Y.: A Secchi depth algorithm considering the residual error in satellite remote sensing reflectance data, Remote Sens., 11(16), 1948., https://doi.org/10.3390/rs11161948, 2019.

Chen, J., Lee, Z. P., Hu, C. M., and Wei, J. W.: Improving satellite data products for open oceans with a scheme to correct the residual errors in remote sensing reflectance, J. Geophys. Res. Oceans, 121, 3866-3886., https://doi.org/10.1002/2016JC011673, 2016.

Chen, J., Quan, W., Jamet, C., and He, X.: Coupling artificial intelligence with a mechanistic model for estimating Secchi depth, Acta Oceanolog. Sin., Accepted., 2025a.

Chen, J. and Quan, W. T.: An improved algorithm for retrieving chlorophyll-a from the Yellow River Estuary using MODIS imagery, Environ. Monit. Assess., 185, 2243-2255., https://doi.org/10.1007/s10661-012-2705-y, 2013.

Chen, J..: A centennial-scale data of suspended particulate matter over the global ocean, Zenodo [data set], https://doi.org/10.5281/zenodo.16991206, 2025b.

Cialdi, M. and Secchi, P. A.: Sur la transparence de la mer, C.R. Hebd. Séances Acad. Sci., 61, 100-104., 1865.

D'Sa, E. J., Miller, R. L., and McKee, B. A.: Suspended particulate matter dynamics in coastal waters from ocean color: Applications to the Northern Gulf of Mexico, Geophys. Res. Lett., 34(23), L23605, https://doi.org/10.1029/2007GL031192, 2007.

Dethier, E. N., Renshaw, C. E., and Magilligan, F. J.: Rapid changes to global river suspended sediment flux by humans, Sci., 376, 1447-1452, https://doi.org/10.1126/science.abn7980, 2022.

Dierssen, H. M.: Perspectives on empirical approaches for ocean color remote sensing of chlorophyll in a changing climate, Proc. Natl. Acad. Sci., 107, 17073-17078, https://doi.org/10.1073/pnas.0913800107, 2010.

https://doi.org/10.5194/essd-2025-484
Preprint. Discussion started: 1 September 2025

Doxaran, D., Froidefond, J. M., and Castaing, P.: A reflectance band ratio used to estimate suspended sediment matter concentrations in sediment-dominated coastal waters, Int. J. Remote Sens., 23(23), 5079-5085., https://doi.org/10.1080/0143116021000009912, 2002.

Doxaran, D., Froidefond, J. M., Castaing, P., and Babin, M.: Dynamics of the turbidity maximum zone in a macrotidal estuary (the Gironde, France): Observations from field and MODIS satellite data, Estuar. Coast. Shelf Sci., 81, 321-337., https://doi.org/10.1016/j.ecss.2008.11.013, 2009.

Effler, S. W., Strait, C., O'Donnell, D. M., Effler, A. J. P., Peng, F., Prestigiacomo, A. R., O'Donnell, S. M., Perkins, M., and Chapra, S. C.: A mechanistic model for secchi disk depth, driven by light scattering constituents, Water Air Soil Poll., 228, 1-24, https://doi.org/10.1007/s11270-017-3323-7, 2017.

Golubkov, M. S. and Golubkov, S. M.: Secchi disk depth or turbidity, which Is better for assessing environmental quality in Eutrophic waters? a case study in a shallow hypereutrophic reservoir, Water, 16, 18, https://doi.org/10.3390/w16010018, 2024.

Gordon, H. R., Clark Dennis, K., Mueller James, L., and Hovis Warren, A.: Phytoplankton Pigments from the Nimbus-7 Coastal Zone Color Scanner: Comparisons with Surface Measurements, Sci., 210, 63-66, https://doi.org/10.1126/science.210.4465.63, 1980a.

Gordon, H. R., Clark, D. K., Mueller, J. L., and Hovis, W. A.: Phytoplankton pigments derived from the Nimbus-7 CZCS: initial comparisons with surface measurements, Sci., 210, 63-66, https://doi.org/10.1126/science.210.4465.63, 1980b.

Green, E. P., Mumby, P. J., Edwards, A. J., and Clark, C. D.: Remote sensing handbook for tropical coastal management, UNESCO, Paris, https://coilink.org/20.500.12592/pg4f7cg, 2000.

He, X. Q., Yan, B., Pan, D. L., Huang, N. L., Xu, D., Chen, J. S., and Cui, Q. F.: Using geostationary satellite ocean color data to map the diurnal dynamics of suspended particulate matter in coastal waters, Remote Sens. Environ., 133, 225-239, https://doi.org/10.1016/j.rse.2013.01.023, 2013.

Hooker, S. B. and Maritorena, S.: An evaluation of oceanographic radiometers and deployment methodologies, Journal of Atmospheric and oceanic technology, 17, 811-830, https://doi.org/10.1175/1520-0426(2000)017<0811:AEOORA>2.0.CO;2, 2000.

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 Goddard Space Flight Center, Greenbelt, Maryland, 1-40 pp., 1994.

Hou, X., Xie, D., Feng, L., Shen, F., and Nienhuis, J. H.: Sustained increase in suspended sediments near global river deltas over the past two decades, Nature Communications, 15, 3319, https://doi.org/10.1038/s41467-024-47598-6, 2024.

IOCCG: Remote sensing of inherent optical properties: fundamentals, tests of algorithms, and applications, Reports of the International Ocean Colour Coordinating Group No.5, IOCCG, Dartmouth, Canada, ISBN 1892926533, 2006.

Jiang, D., Matsushita, B., Pahlevan, N., Gurlin, D., Lehmann, M. K., Fichot, C. G., Schalles, J., Loisel, H., Binding, C., Zhang, Y., Alikas, K., Kangro, K., Uusõue, M., Ondrusek, M., Greb, S., Moses, W. J., Lohrenz, S., and O'Donnell, D.:

Remotely estimating total suspended solids concentration in clear to extremely turbid waters using a novel semi-analytical method, Remote Sens. Environ., 258, 112386, https://doi.org/10.1016/j.rse.2021.112386, 2021.

Kim, J., Kwon, S., Chung, S., and Kim, Y. D.: Turbidity and suspended sediment relationship based on sediment composition and particle size distribution, Sci. Rep., 15, 16286, https://doi.org/10.1038/s41598-025-00435-2, 2025.

Lavigne, H., Dogliotti, A., Doxaran, D., Shen, F., Castagna, A., Beck, M., Vanhellemont, Q., Sun, X., Gossn, J. I., Renosh, P. R., Sabbe, K., Vansteenwegen, D., and Ruddick, K.: The HYPERMAQ dataset: bio-optical properties of moderately to extremely turbid waters, Earth Syst. Sci. Data, 14, 4935-4947, https://doi.org/10.5194/essd-14-4935-2022, 2022.

Loisel, H., Mangin, A., Vantrepotte, V., Dessailly, D., Ngoc Dinh, D., Garnesson, P., Ouillon, S., Lefebvre, J.-P., Mériaux, X., and Minh Phan, T.: Variability of suspended particulate matter concentration in coastal waters under the Mekong's influence from ocean color (MERIS) remote sensing over the last decade, Remote Sens. Environ., 150, 218-230, https://doi.org/10.1016/j.rse.2014.05.006, 2014.

Mao, Z., Chen, J., Pan, D., Tao, B., and Zhu, Q.: A regional remote sensing algorithm for total suspended matter in the East China Sea, Remote Sens. Environ., 124, 819–831, https://doi.org/10.1016/j.rse.2012.06.014, 2012.

Milliman, J. D. and Meade, R. H.: World-wide delivery of fiver sediment to the oceans, J. Geol., 91, 1-21, 1983.

Morel, A. and Prieur, L.: Analysis of variances in ocean color, Limnol. Oceanogr. 22, 709-722, https://doi.org/10.1086/628741, 1977.

Mueller, J. L., Fargion, G. S., and McClain, C. R.: Ocean optics protocols for satellite ocean color sensor validation, revision 4, Goddard Space Flight Center, Greenbelt, MD, 84 pp., https://doi.org/10.25607/OBP-62, 2003.

Nations, U.: Department of Economic and Social Affairs, Population Division, United Nations, 2022.

Nechad, B., Ruddick, K. G., and Park, Y.: Calibration and validation of a generic multisensor algorithm for mapping of total suspended matter in turbid waters, Remote Sens. Environ., 114, 854-866, https://doi.org/10.1016/j.rse.2009.11.022, 2010.

Overeem, I., Hudson, B. D., Syvitski, J. P. M., Mikkelsen, A. B., Hasholt, B., van den Broeke, M. R., Noël, B. P. Y., and Morlighem, M.: Substantial export of suspended sediment to the global oceans from glacial erosion in Greenland, Nat Geosci., 10, 859-863, https://doi.org/10.1038/ngeo3046, 2017.

Petus, C., Chust, G., Gohin, F., Doxaran, D., Froidefond, J. M., and Sagarminaga, Y.: Estimating turbidity and total suspended matter in the Adour River plume (South Bay of Biscay) using MODIS 250-m imagery, Cont. Shelf Res., 30, 379-392, https://doi.org/10.1016/j.csr.2009.12.007, 2010.

Pitarch, J.: A review of Secchi's contribution to marine optics and the foundation of Secchi disk science, Oceanography, 33, 26-37, https://www.jstor.org/stable/26962479, 2020.

Preisendorfer, R. W.: Secchi disk science: Visual optics of natural waters, Limnol. Oceanogr., 31, 909-926, https://doi.org/10.4319/lo.1986.31.5.0909, 1986.

Rodriguez, A. B., McKee, B. A., Miller, C. B., Bost, M. C., and Atencio, A. N.: Coastal sedimentation across North America doubled in the 20th century despite river dams, Nat. Commun., 11, 3249, https://doi.org/10.1038/s41467-020-16994-z, 2020.

https://doi.org/10.5194/essd-2025-484
Preprint. Discussion started: 1 September 2025

Scown, M. W., Dunn, F. E., Dekker, S. C., van Vuuren, D. P., Karabil, S., Sutanudjaja, E. H., Santos, M. J., Minderhoud, P. S. J., Garmestani, A. S., and Middelkoop, H.: Global change scenarios in coastal river deltas and their sustainable development implications, Glob. Environ. Change, 82, 1-14, https://doi.org/10.1016/j.gloenvcha.2023.102736, 2023.

Shen, F. and Verhoef, W.: Suppression of local haze variations in MERIS images over turbid coastal waters for retrieval of suspended sediment concentration, Opt. Express., 18, 12653-12662, https://doi.org/10.1364/OE.18.012653, 2010.

Sokoletsky, L., Fang, S., Yang, X., and Wei, X.: Evaluation of empirical and semianalytical spectral reflectance models for surface suspended sediment concentration in the highly variable estuarine and coastal waters of east China, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., 9, 5182-5192, https://doi.org/10.1109/JSTARS.2016.2582909, 2016.

Strakhovenko, V. D., Ovdina, E. A., Malov, V. I., and Malov, G. I.: Chemical elements concentration, variations in mineral composition, and current rate of sedimentation in sapropel deposits of small lakes in the taiga zone of southern Siberia. In: Russian Geology and Geophysics, Novosibirsk State University, 64, 824–838, https://doi.org/10.2113/RGG20234526, 2023.

Strickland, J. D. H. and Parsons, T. R.: A practical handbook of seawater analysis (2nd edition), Fisheries Research Board of Canada Bulletin, No. 167, Fisheries Research Board of Canada, Ottawa, Canada, 310 pp., ISBN 0660115963, 1972.

Sun, D., Li, Y., Le, C., Shi, K., Huang, C., Gong, S., and Yin, B.: A semi-analytical approach for detecting suspended particulate composition in complex turbid inland waters (China), Remote Sens. Environ., 134, 92-99, https://doi.org/10.1016/j.rse.2013.02.024, 2013.

Swift, T. J., Perez-Losada, J., Schladow, S. G., Reuter, J. E., Jassby, A. D., and Goldman, C. R.: Water clarity modeling in Lake Tahoe: Linking suspended matter characteristics to Secchi depth, Aquatic Sciences, 68, 1-15, https://doi.org/10.1007/s00027-005-0798-x, 2006.

Szeligowska, M., Benkort, D., Przyborska, A., Moskalik, M., Moreno, B., Trudnowska, E., and Błachowiak-Samołyk, K.: Estimates of carbon sequestration potential in an expanding Arctic fjord (Hornsund, Svalbard) affected by dark plumes of glacial meltwater, BG, 21, 3617-3639, https://doi.org/10.5194/bg-21-3617-2024, 2024.

Teng, W., Yu, Q., Stramski, D., Reynolds, R. A., Woodruff, J. D., and Yellen, B.: High spatial-resolution satellite mapping of suspended particulate matter in global coastal waters using particle composition-adaptive algorithms, Remote Sens. Environ., 323, 114745, https://doi.org/10.1016/j.rse.2025.114745, 2025.

Tyler, J. E.: The Secchi disc, Limnol. Oceanogr., 13, 1-6, https://doi.org/10.4319/lo.1968.13.1.0001, 1968.

Valente, A., Sathyendranath, S., Brotas, V., Groom, S., Grant, M., Jackson, T., Chuprin, A., Taberner, M., Airs, R., Antoine, D., Arnone, R., Balch, W. M., Barker, K., Barlow, R., Bélanger, S., Berthon, J. F., Beşiktepe, Ş., Borsheim, Y., Bracher, A., Brando, V., Brewin, R. J. W., Canuti, E., Chavez, F. P., Cianca, A., Claustre, H., Clementson, L., Crout, R., Ferreira, A., Freeman, S., Frouin, R., García-Soto, C., Gibb, S. W., Goericke, R., Gould, R., Guillocheau, N., Hooker, S. B., Hu, C., Kahru, M., Kampel, M., Klein, H., Kratzer, S., Kudela, R., Ledesma, J., Lohrenz, S., Loisel, H., Mannino, A., Martinez-Vicente, V., Matrai, P., McKee, D., Mitchell, B. G., Moisan, T., Montes, E., Muller-Karger, F., Neeley, A., Novak, M., O'Dowd, L., Ondrusek, M., Platt, T., Poulton, A. J., Repecaud, M., Röttgers, R., Schroeder, T., Smyth, T., Smythe-Wright, D., Sosik, H. M., Thomas, C., Thomas, R., Tilstone, G., Tracana, A., Twardowski, M., Vellucci, V., Voss, K., Werdell, J.,

https://doi.org/10.5194/essd-2025-484
Preprint. Discussion started: 1 September 2025

Wernand, M., Wojtasiewicz, B., Wright, S., and Zibordi, G.: A compilation of global bio-optical in situ data for ocean colour satellite applications – version three, Earth Syst. Sci. Data, 14, 5737-5770, https://doi.org/10.5194/essd-14-5737-2022, 2022.

Werdell, P. J. and Bailey, S. W.: The SeaWiFS bio-optical archive and storage system (SeaBASS): current architecture and implementation, Goddard Space Flight Center, Greenbelt, Maryland 20771., 2002. 2002.

Wozniak, S. B., Stramski, D., Stramska, M., Reynolds, R. A., Wright, V. M., Miksic, E. Y., Cichocka, M., and Cieplak, A. M.: Optical variability of seawater in relation to particle concentration, composition, and size distribution in the nearshore marine environment at Imperial Beach, California, J. Geophys. Res. Oceans, 115, -, https://doi.org/10.1029/2009JC005554, 2010.

Yoshida, A., Tobo, Y., Adachi, K., Moteki, N., Kawai, Y., Sasaoka, K., and Koike, M.: Analysis of oceanic suspended particulate matter in the western North Pacific using the complex amplitude sensor, Sci. Rep., 14, 20055, https://doi.org/10.1038/s41598-024-70683-1, 2024.

https://doi.org/10.5194/essd-2025-484
Preprint. Discussion started: 1 September 2025

Table 1: Data sources and used

Source	Temporal	Parameter
https://www.ncei.noaa.gov	1991-2014	Z_{sd}
https://www.bodc.ac.uk	2013-2018	Z_{sd}
https://www.jodc.go.jp	1986	Z_{sd}
https://ecos.fws.gov	1899-2019	Z_{sd}
https://data.ontario.ca	2003-2005	Z_{sd}
https://www.data.gov.in	2007-2021	Z_{sd}
https://wesr.unep.org	2015	Z_{sd}
https://doi.pangaea.de/10.1594/	1903-2020	Z_{sd}
https://seabass.gsfc.nasa.gov/	1979-2020	Z_{sd} and R_{rs}
Boyce et al. (2010)	1890-2010	Z_{sd}
Chen et al. (2015; 2019)	2003-2018	Z_{sd}, R_{rs}, and SPM
Lavigne et al. (2022)	2018	R_{rs} and SPM
Valente et al. (2022)	1997-2021	R_{rs}
https://www.pangaea.de	1991-2019	SPM
https://www.waterqualitydata.us	1966-2025	SPM

475

https://doi.org/10.5194/essd-2025-484
Preprint. Discussion started: 1 September 2025

Open Access Earth System Science Data Discussions

Table 2: Descriptive statistics of bio-optical dataset used for model initialization and assessment. Std represents the standard deviation

Model	Variable	Min	Max	Median	Average	Std
SPM_{zsd}	Calibration dataset (*N*=217)					
	Z_{sd}	0.18	21.00	4.50	5.69	4.47
	SPM	0.50	120.48	2.88	7.64	14.20
	Latitude	20.19	39.55	33.50	31.77	5.43
	Longtitude	109.04	125.01	122.01	120.15	4.86
	Validation dataset (*N*=148)					
	Z_{sd}	1.20	45.00	4.70	14.13	13.79
	SPM	0.13	16.8	3.60	4.89	4.68
	Latitude	16.21	39.95	29.97	29.21	10.22
	Longtitude	108.31	129.40	119.38	116.23	5.24
SPM_{rrs}	Calibration dataset (*N*=620)					
	X	-0.18	1.25	0.38	0.47	0.35
	SPM	0.10	2239.80	12.77	109.25	296.69
	Latitude	-35.58	51.59	32.69	32.57	11.50
	Longtitude	-70.34	125.68	120.85	93.98	59.06
	Validation dataset (*N*=198)					
	X	-0.27	1.13	0.30	0.30	0.41
	SPM	0.13	577.2	7.95	42.32	84.61
	Latitude	16.86	39.91	31.25	29.55	8.79
	Longtitude	108.31	122.74	118.77	116.89	4.63

480

https://doi.org/10.5194/essd-2025-484
Preprint. Discussion started: 1 September 2025

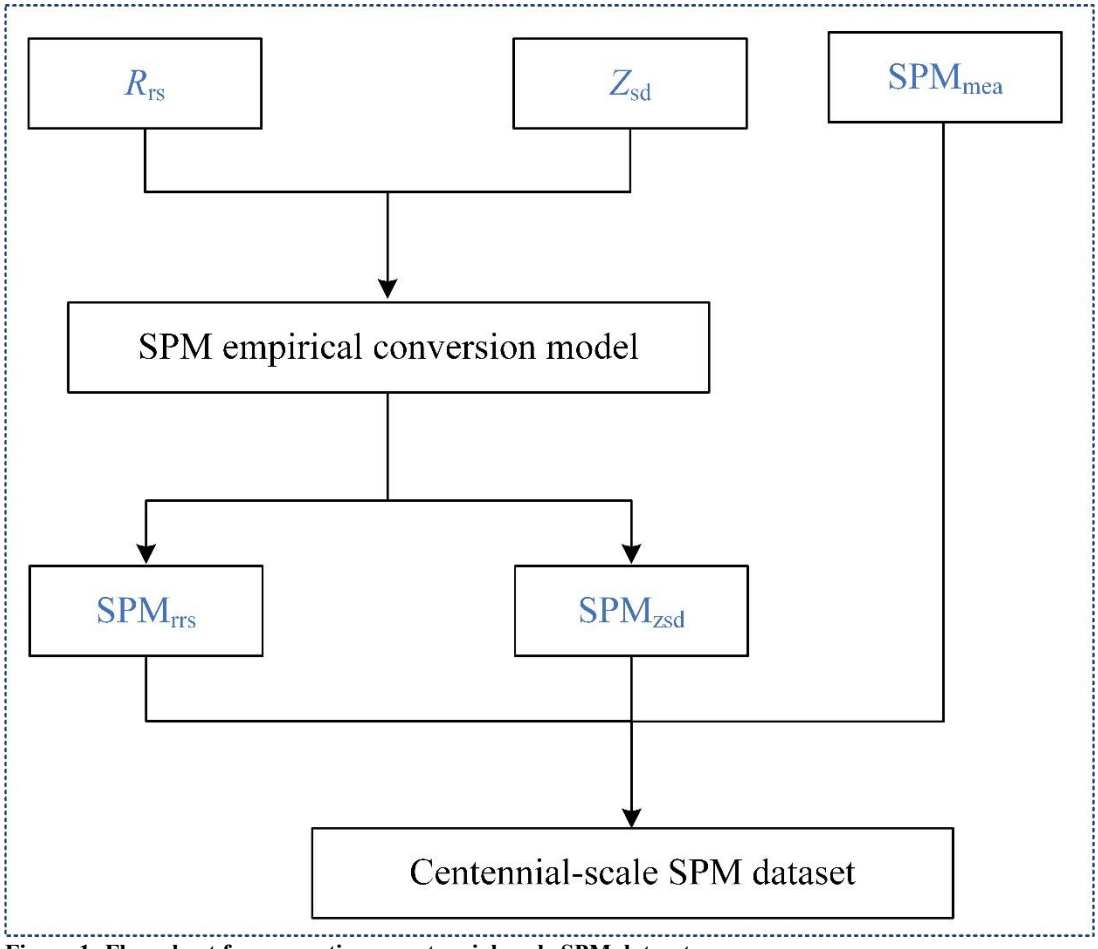

Figure 1: Flow chart for generating a centennial-scale SPM dataset

https://doi.org/10.5194/essd-2025-484
Preprint. Discussion started: 1 September 2025

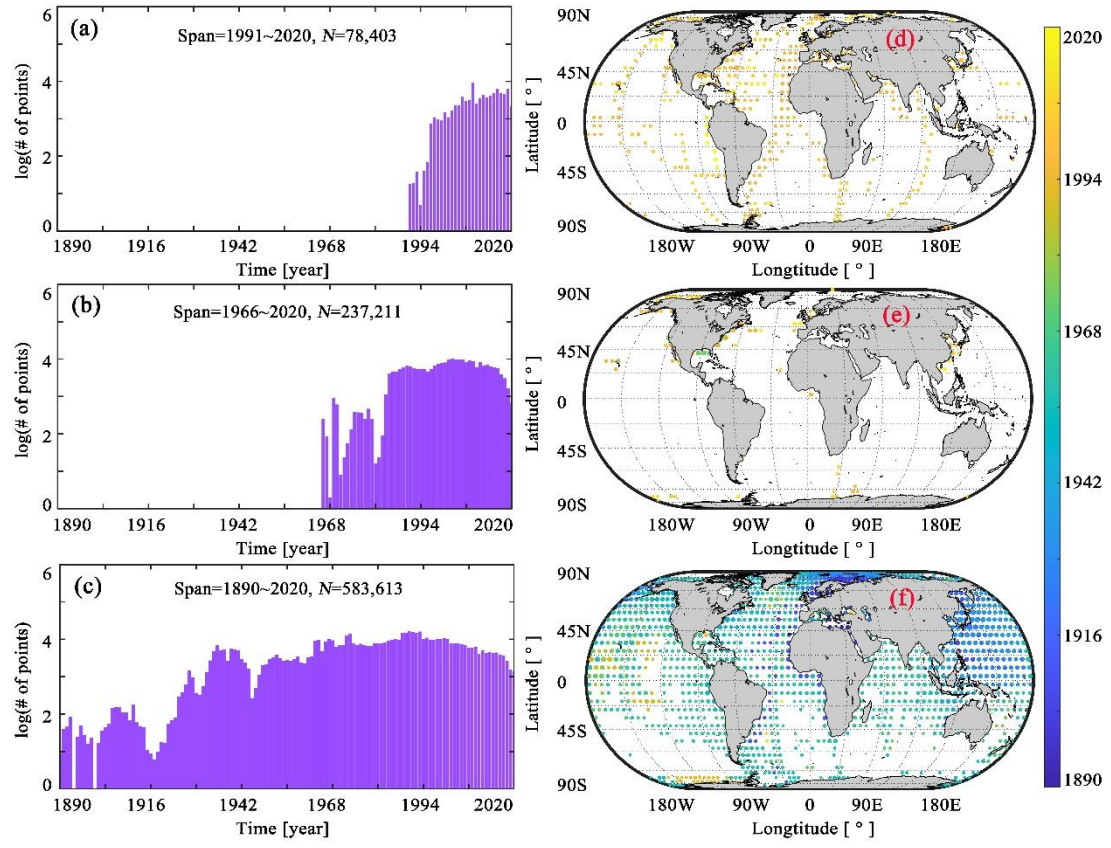

Figure 2: Global oceanic data availability. Temporal coverage of (a–b) bottle-derived SPM measurements, (c–d) remote sensing reflectance observations, and (e–f) ocean transparency records. The global ocean is partitioned into 5° × 5° grid cells. For each panel group, the first column depicts the fraction of total observations acquired per year, the second column indicates the year of the earliest available record within each grid cell, and the third column presents the baseline year together with the full temporal span of data availability in each cell.

https://doi.org/10.5194/essd-2025-484
Preprint. Discussion started: 1 September 2025

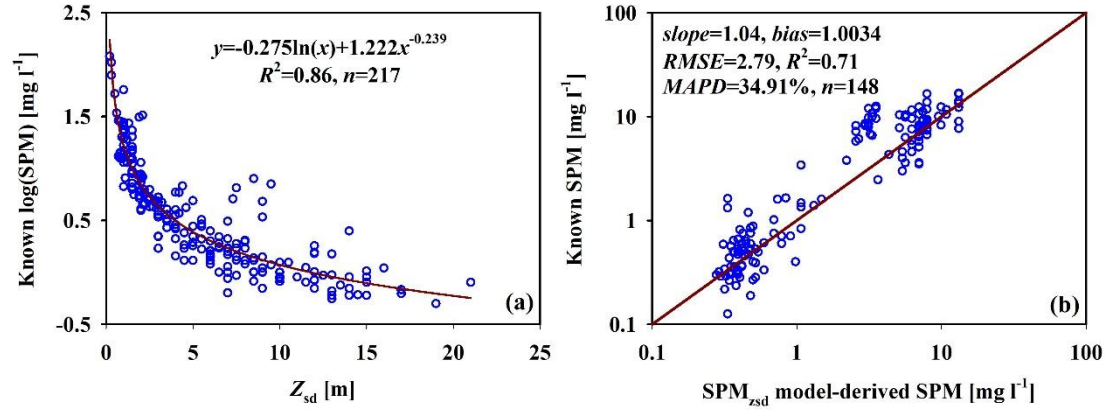

Figure 3: Empirical model for deriving SPM from transparency measurements: (a) model initialization and (b) model evaluation.

https://doi.org/10.5194/essd-2025-484
Preprint. Discussion started: 1 September 2025

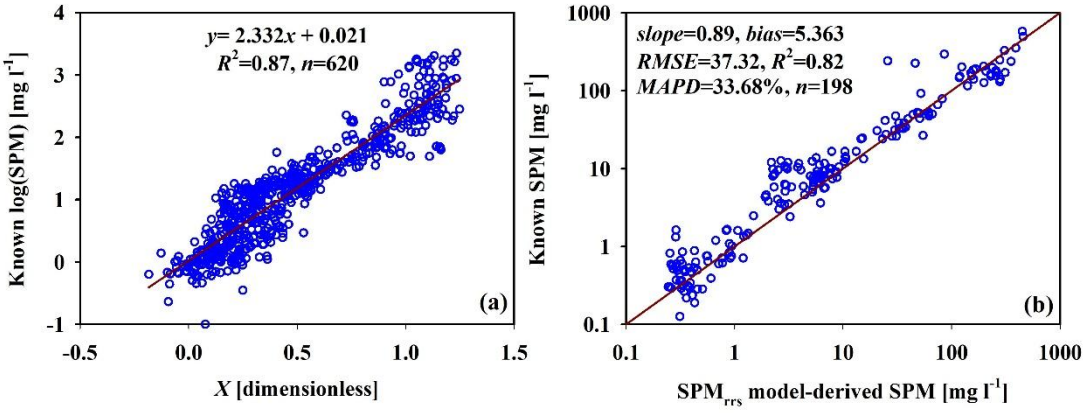

Figure 4: Empirical model for deriving SPM from remote sensing reflectance measurements: (a) model initialization and (b) model evaluation.

https://doi.org/10.5194/essd-2025-484
Preprint. Discussion started: 1 September 2025

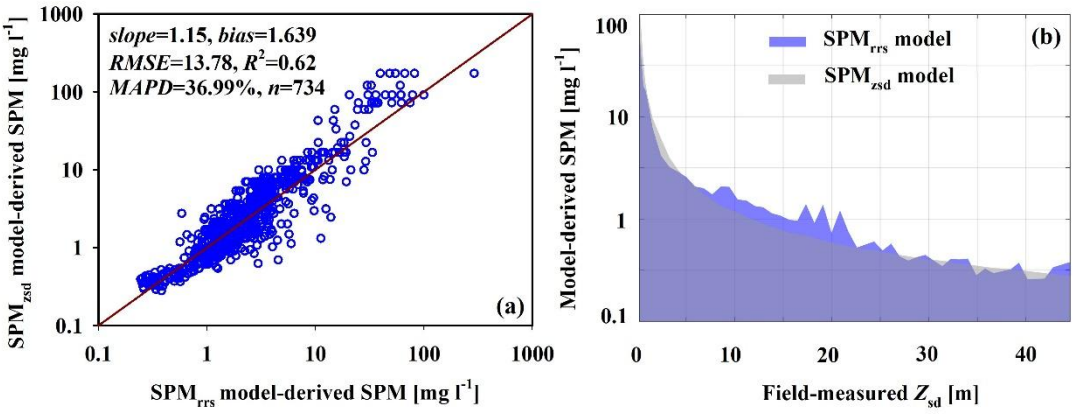

Figure 5: Consistency analysis of SPM data from different sources. (a) Scatterplot comparing SPM derived from transparency measurements with SPM retrieved from remote sensing reflectance measurements; (b) histogram of differences between SPM derived from transparency measurements and SPM retrieved from remote sensing reflectance measurements.

https://doi.org/10.5194/essd-2025-484
Preprint. Discussion started: 1 September 2025

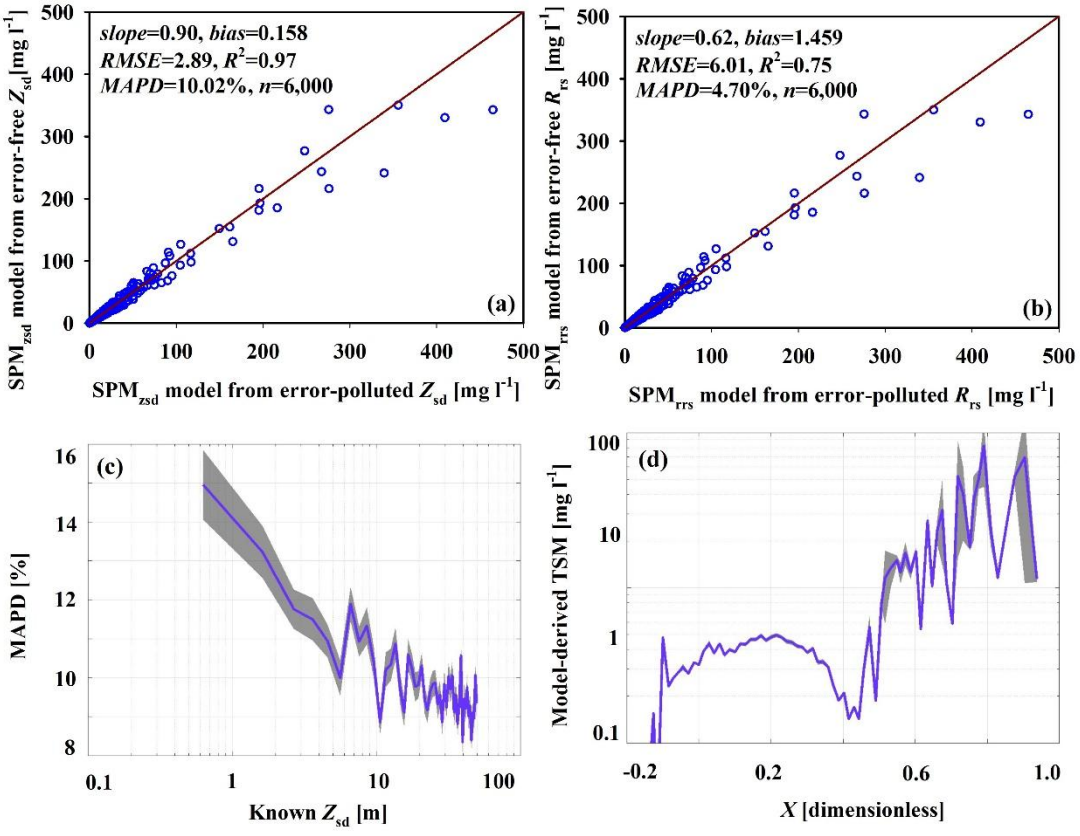

Figure 6: Sensitivity analysis of empirical SPM retrieval models to data errors by comparing SPM derived from noise-free and noise-contaminated data. (a) SPM estimates with and without artificially added noise to Zsd data;(c) corresponding MAPD histograms;(b) SPM estimates with and without artificially added noise to Rrs data;(d) corresponding MAPD histograms. Shaded areas in the histograms represent variance ranges.

https://doi.org/10.5194/essd-2025-484
Preprint. Discussion started: 1 September 2025

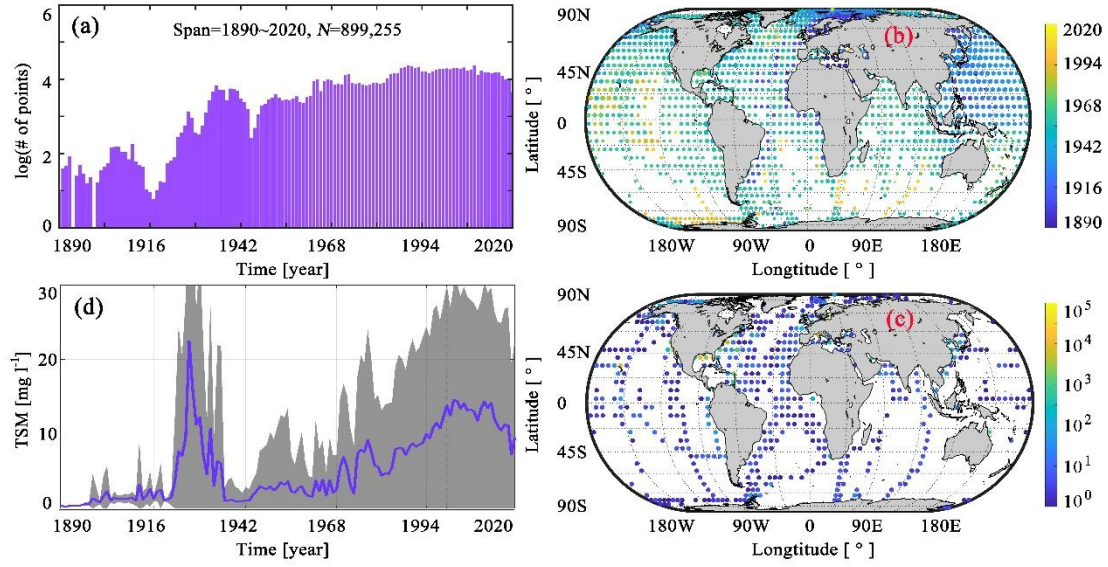

Figure 7: Centennial-scale dataset of SPM over the global ocean derived from multi-source data fusion.(a) Temporal coverage of total suspended sediment concentration; (b) Acquisition time of the earliest available measurement within each 5° × 5° grid cell for the corresponding datasets; (c) Increase in global grid sample density compared with SPM derived from Zsd data within each 5° × 5° grid cell; (d) Interannual variability of global ocean SPM over the centennial scale derived from the fused dataset. Bars indicate the logarithmic number of measurements collected annually.