INTERNATIONAL JOURNAL OF
COASTAL, OFFSHORE & ENVIRONMENTAL ENGINEERING

Available online at: www. ijcoe.org

Assessing Satellite–In Situ Collocation Uncertainty in Sentinel-2–Based Coastal Turbidity Monitoring

Amirmahdi Zarboo[1*]

[1*] *Iranian National Institute for Oceanography and Atmospheric Sciences (INIOAS), Tehran, 1411813389, Iran*
amirmahdizarboo@gmail.com

ARTICLE INFO

Article History:
Received….
Accepted ….
Available online ….

Keywords:
Sentinel-2
Coastal turbidity
Satellite–in situ collocation
Water quality
Operational monitoring
Reflectance indices

ABSTRACT

Accurate satellite–in-situ collocation remains a key source of uncertainty in quantitative retrievals of coastal water turbidity, particularly in shallow and optically complex environments. This study presents a systematic assessment of how spatial, temporal, and depth-related collocation choices influence the performance of commonly used Sentinel–2 reflectance-based indices for turbidity monitoring. Three indices—the Normalized Difference Turbidity Index (NDTI), the Normalized Difference Water Index (NDWI), and the Normalized Difference Vegetation Index (NDVI)—were evaluated against an extensive in-situ dataset of CTD turbidity measurements aggregated into three near-surface depth layers (0–1 m, 1–2 m, and 2–3 m). Satellite–in-situ matching was conducted using spatial buffers ranging from 120 to 400 m, two processing resolutions (20 and 40 m), and a temporal matching window of up to ±48 h. Across all collocation scenarios, NDWI exhibited the strongest and most stable relationship with turbidity (FTU), with regression slopes and explained variance remaining consistent across depth layers, buffer sizes, and spatial resolutions. NDVI showed a weaker but coherent negative response, while NDTI demonstrated negligible predictive skill within the observed turbidity range. The limited sensitivity of retrieval performance to collocation parameter choices indicates that, for moderate turbidity conditions in shallow coastal waters, robust satellite-based estimates can be achieved without highly restrictive matching criteria. The proposed workflow provides a transferable and reproducible framework for evaluating satellite–in-situ collocation uncertainty and supports routine and operational coastal water-quality monitoring using Sentinel–2 data.

1. Introduction

Accurate monitoring of coastal water turbidity is essential for understanding sediment transport, water quality, and benthic habitat conditions in shallow and optically complex environments. Satellite remote sensing, and in particular Sentinel–2 multispectral imagery, provides spatially and temporally consistent observations that support routine coastal water-quality assessments at regional to local scales. However, quantitative retrievals of turbidity from satellite reflectance remain strongly influenced by uncertainties associated with satellite–in-situ collocation, including spatial heterogeneity, temporal mismatch between acquisitions and field measurements, and vertical variability within the near-surface water column. These sources of uncertainty complicate the interpretation and transferability of empirical relationships between satellite-derived indices and in-situ turbidity observations.

A wide range of reflectance-based indices has been proposed to estimate coastal turbidity, including the Normalized Difference Turbidity Index (NDTI), the Normalized Difference Water Index (NDWI), and the Normalized Difference Vegetation Index (NDVI). While these indices have been evaluated across diverse coastal and inland water bodies, their reported performance often varies substantially between studies, partly due to differences in satellite–in-situ matching strategies. In practice, choices related to spatial buffer size, processing resolution, temporal matching windows, and depth aggregation of in-situ measurements are frequently made heuristically and differ widely across applications. As a result, it remains

unclear to what extent observed differences in index performance reflect intrinsic optical sensitivity versus artifacts of collocation methodology. Addressing this gap requires systematic and reproducible assessments of collocation parameter sensitivity under realistic operational conditions.

Coastal waters are dynamic environments where suspended particulate matter (SPM), colored dissolved organic matter (CDOM), sediment resuspension, and riverine inputs jointly influence water transparency, optical properties, and ecosystem functioning. Turbidity, a proxy for SPM concentration, plays a key role in controlling light penetration, photosynthetic availability, and the health of benthic habitats such as seagrass meadows [1,2]. Accurate characterization of turbidity is therefore essential for understanding coastal biogeochemistry and for informing management activities aimed at mitigating anthropogenic pressures [3].

Remote sensing provides an effective means of monitoring turbidity across spatial and temporal scales that exceed the capabilities of in-situ sampling. Theoretical and empirical studies demonstrate that variations in water reflectance, particularly in the green and red–NIR spectral regions, are highly sensitive to SPM concentrations [4–6]. The launch of Sentinel–2 has been particularly transformative, offering 10–20 m spatial resolution, a five-day revisit cycle, and spectral bands optimized for coastal water-quality monitoring [7,8]. As a result, Sentinel–2 products have been successfully applied to derive turbidity [e.g., 9,10], detect submerged vegetation, map sediment plumes, and support operational coastal observing systems.

Despite these advances, several challenges remain. First, relationships between satellite-derived spectral indices and turbidity can vary with viewing conditions, water depth, substrate type, and atmospheric correction uncertainties [11]. Second, optical gradients in shallow environments may be influenced by benthic reflectance, complicating the interpretation of near-surface reflectance signals [12,13]. Third, the spatial scale of satellite pixels may not always correspond to that of in-situ measurements, requiring methodological choices regarding spatial averaging and buffer size [14]. Finally, temporal mismatch between satellite overpasses and field measurements can lead to decorrelation, particularly in highly dynamic coastal systems [15].

To address these issues, recent studies have explored the use of simple reflectance-based indices derived from Sentinel–2, such as the Normalized Difference Water Index (NDWI), the Normalized Difference Vegetation Index (NDVI), and the Normalized Difference Turbidity Index (NDTI), as proxies for SPM and turbidity. NDWI, originally proposed to delineate open water, has proven sensitive to turbidity due to contrasting green and NIR reflectance behavior [16,17]. NDVI, although traditionally used for vegetation mapping, decreases with increasing turbidity because red–NIR contrast diminishes in optically dense waters [18]. NDTI, defined from red and green reflectance, has been applied to discriminate sediment-laden waters in several coastal studies [9,17]. However, the relative performance and robustness of these indices under different spatial averaging schemes, depth structures, and turbidity ranges remain insufficiently understood.

In this context, the present study provides a rigorous evaluation of Sentinel–2 reflectance indices for quantifying turbidity in a shallow coastal environment using an extensive set of depth-resolved CTD measurements. I combine depth-aggregated turbidity profiles with a flexible collocation framework that accounts for spatial buffer radius, pixel resolution, and temporal matching. Our objectives are threefold: (1) to assess the sensitivity of NDWI, NDVI, and NDTI to measured turbidity; (2) to determine the influence of depth structure and spatial averaging on index performance; and (3) to quantify the robustness of index–turbidity relationships across a range of realistic parameter configurations. By resolving these methodological uncertainties, the study provides practical guidance for operational coastal monitoring and contributes to the advancement of semi-empirical turbidity retrieval from multispectral satellite imagery.

2. Materials and Methods

2.1. Study Area

The study area is situated along the shallow coastline of the Persian Gulf, where seagrass meadows are present and water depths rarely exceed one meter at low tide. Capability for frequent sediment resuspension due to tidal currents and wind-driven waves renders this zone highly dynamic in turbidity, challenging remote sensing retrievals near the seabed and under variable bottom reflectance conditions [19]. Seagrass ecosystems in such environments are especially sensitive to elevated turbidity and reduced light penetration [20].

The distribution of CTD sampling stations across the study region is shown in Figure 1. The vertical structure of turbidity in the upper 10 m of the water column is summarized in Figure 2.

2.2 CTD Dataset and Preprocessing

A vessel-based survey collected a total of 63,166 CTD casts during the period 7 September 2021 to 30 September 2021, each including turbidity, depth, temperature and salinity. To eliminate bottom-influenced measurements, I retained only observations with depth ≤ 1 m — a threshold chosen to ensure minimal bottom reflectance influence [21]. All rows with missing geographic coordinates, timestamps, or turbidity values were excluded. Because the vessel revisited near-identical positions, data were spatially aggregated by unique (latitude, longitude) pairs using both mean and median statistics depending on the scenario. All timestamps were converted to UTC to

align with satellite overpass times. After filtering, the final dataset comprised 2,822 shallow-water points. Turbidity was measured with a Seapoint turbidity meter integrated in a CTD75M probe (Sea & Sun Technology, Germany). The sensor is factory-calibrated against formazin standards and outputs turbidity in Formazin Turbidity Units (FTU). For formazin-based standards, FTU are numerically equivalent to the more commonly used Nephelometric Turbidity Units (NTU), and I therefore report turbidity in FTU.

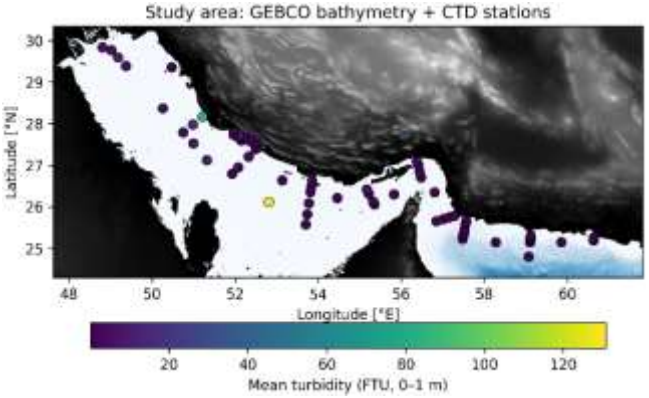

Figure 1 Study area showing GEBCO bathymetry (water depth shaded in blue and land elevation in gray) together with the distribution of CTD measurement stations used in this study. Station markers represent near-surface turbidity (0–1 m), averaged at each unique sampling location. The map highlights the spatial variability of the shallow coastal environment, including the narrow shelf region where Sentinel–2 observations were collocated with in-situ turbidity measurements.

2.3. Sentinel-2 Data Acquisition and Preprocessing

Surface reflectance products from the Sentinel-2 program (Level-2A) were retrieved via the Sentinel-Hub "classic" API. For each CTD point I queried scenes within temporal windows of ±6 h and ±120 h, and applied two cloud-coverage thresholds (10% and 100%). To account for vessel-drift and pixel footprint, circular buffers of 200 m and 300 m radius were generated around each CTD location. A custom evaluation script computed three indices:

$$\mathrm{NDWI} = \frac{\mathrm{B3}-\mathrm{B8}}{\mathrm{B3}+\mathrm{B8}} \quad (1)$$

$$NDVI = \frac{B8-B4}{B8+B4} \quad (2)$$

$$NDTI = \frac{B11-B12}{B11+B12} \quad (3)$$

2.4. Collocation Framework

For each valid CTD measurement the collocation workflow proceeded as follows: Sentinel-2 scenes satisfying the temporal and cloud criteria were identified, pixels within the spatial buffer were extracted, and the masks described in Section 2.3 were applied. The remaining pixel values were aggregated either by mean or median to produce a single index value per CTD point. Metadata—such as the number of valid pixels (n), temporal difference between CTD and overpass, cloud coverage and scene ID—were logged. Points with zero valid pixels after masking were excluded from further analysis [22].

The overall processing chain from raw CTD profiles to depth-resolved Sentinel–2 collocations is summarized in Figure 3.

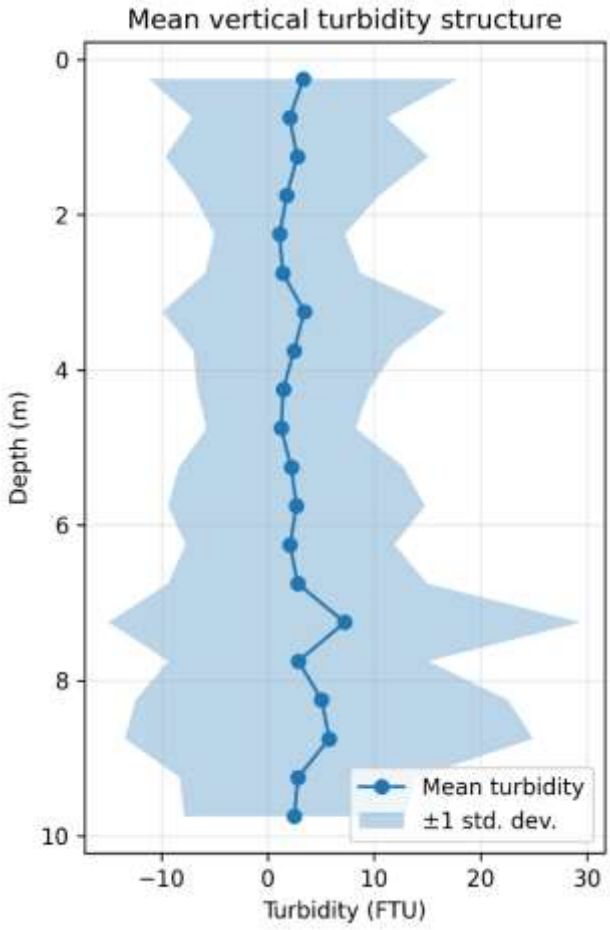

Figure 2 Mean vertical turbidity structure derived from all CTD profiles in the survey. Turbidity (FTU) is averaged in 0.5 m depth bins up to 10 m depth, with the shaded envelope indicating ±1 standard deviation. The absence of strong gradients within the upper 3 m suggests that near-surface optical signals from Sentinel–2 are representative of the shallow water column, supporting the use of depth-aggregated turbidity layers (0–1 m, 1–2 m, 2–3 m) for satellite collocation.

2.5. Depth-resolved relationships between turbidity and Sentinel-2 indices

To reduce noise from repeated casts at the same station and to explore potential vertical structure in the turbidity signal, the CTD measurements were first aggregated by depth layer and position. For each CTD profile, turbidity values were averaged within three discrete depth intervals (0–1 m, 1–2 m, 2–3 m) at each longitude/latitude station. Here I focus on the 1–2 m and 2–3 m layers, for which collocations with Sentinel-2 were successful and yielded consistent sample sizes of n = 53 stations per layer. For each depth-aggregated CTD dataset, I performed independent collocations with three Sentinel-2 band-derived indices: the Normalized Difference Turbidity Index (NDTI), the Normalized Difference Water Index (NDWI), and the Normalized Difference Vegetation Index (NDVI). All

collocations used Sentinel-2 L2A data, a temporal matching window of ±48 h with a maximum scene cloud fraction of 40 %, spatial buffers of 120–400 m around each CTD station, and processing resolutions of 20 m and 40 m. Across this parameter space, the strength and sign of the index–turbidity relationships were very similar, indicating that the results are robust to moderate changes in spatial buffer and pixel size. For brevity, Table 1 summarizes the regression statistics for a representative configuration (200 m buffer, 20 m resolution).

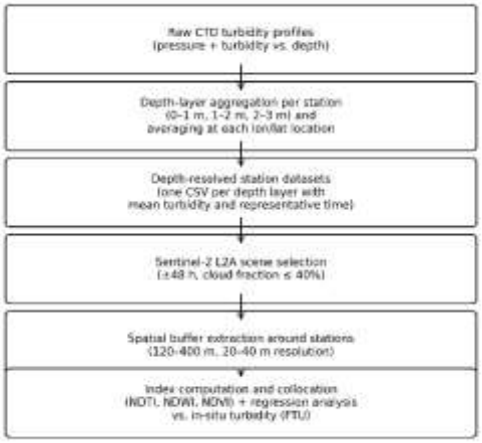

Figure 3 Overview of the CTD–Sentinel–2 collocation workflow used in this study. CTD turbidity profiles are first aggregated into three discrete depth layers (0–1 m, 1–2 m, 2–3 m) and spatially grouped by unique longitude/latitude coordinates. Each station is assigned a representative sampling time based on the median measurement time within the layer. Sentinel–2 L2A scenes are then selected using a ±48 h temporal window and a maximum cloud fraction of 40 %. For each station, spatial buffers of 120–400 m are extracted at 20 m or 40 m resolution, and three indices (NDTI, NDWI, NDVI) are derived. The resulting depth-resolved collocation dataset provides a consistent basis for quantifying index sensitivity to near-surface turbidity.

NDWI showed the clearest and most consistent relationship with in situ turbidity (FTU) in both depth layers. At 1–2 m, NDWI explained approximately 13 % of the variance in turbidity ($R^2 = 0.134$; n = 53), with a positive regression slope of ∼ 34 FTU per unit NDWI. The relationship was slightly stronger for the 2–3 m layer ($R^2 = 0.141$; n = 53), with a very similar slope (∼ 35 FTU per unit NDWI). In other words, stations with higher NDWI values tend to coincide with more turbid water, and this tendency is stable across the upper 3 m of the water column.

NDVI also exhibited a systematic response to turbidity, but with an opposite sign and slightly lower explanatory power. For the 1–2 m layer, NDVI and turbidity were negatively correlated, with $R^2 = 0.114$ (n = 53) and a slope of about −65 FTU per unit NDVI. The 2–3 m layer showed a comparable pattern ($R^2 = 0.121$; slope ≈−67 FTU per unit NDVI). These negative slopes are consistent with the expectation that increasing turbidity reduces the relative contribution of benthic or submerged vegetation to the water-leaving signal, leading to lower NDVI values.

By contrast, NDTI showed only very weak relationships with turbidity in all configurations and depth layers. In the 1–2 m and 2–3 m layers, R^2 remained below 0.01 (Table 1), indicating that the SWIR-based NDTI is essentially insensitive to the range of turbidity conditions observed during the survey. This confirms that, for the present study area and period, NDTI is not a useful predictor of FTU-based turbidity, whereas NDWI and NDVI carry a modest but consistent turbidity signal.

Overall, these results demonstrate that simple band-ratio indices derived from Sentinel- 2 (in particular NDWI, and to a lesser extent NDVI) retain a monotonic and depth-robust relationship with in situ turbidity in the upper 3 m of the water column. Although the explained variance is moderate (on the order of 10–14%), the relationships are remarkably stable across depth layers, spatial buffers and processing resolutions, providing a useful basis for subsequent development of semi-empirical turbidity retrievals or for relative turbidity mapping in the study area.

Table 1 Summary of linear relationships between Sentinel-2 indices and in-situ turbidity (FTU) for different depth layers. Results shown for the configuration with 200 m spatial buffer and 20 m pixel resolution.

Index	Depth layer	n	R^2	Slope (FTU per unit index)
NDTI	1–2 m	53	0.006	39.1
NDVI	1–2 m	53	0.114	-65.4
NDWI	1–2 m	53	0.134	34.3
NDTI	2–3 m	53	0.008	43.1
NDVI	2–3 m	53	0.121	-66.8
NDWI	2–3 m	53	0.141	34.8

2.6. Parameter-Sensitivity Scenarios

To quantify the sensitivity of turbidity–index relationships to processing decisions, I evaluated all combinations of the following parameters: buffer radius (200 m vs. 300 m), time window (±6 h vs. ±120 h), cloud threshold (10% vs. 100%), water-only mask (enabled vs. disabled) and aggregation statistic (mean vs. median). Each unique configuration produced a collocation dataset, enabling comparative analysis of regression stability [23]. Although all combinations of parameters were tested during exploratory analysis, four representative scenarios were retained for detailed evaluation to ensure clarity and interpretability (Table 2). These scenarios isolate the effects of spatial buffering, temporal matching, and water-only masking, while maintaining a consistent baseline configuration.

2.7. Statistical Analysis

For each scenario and each spectral index, scatterplots of index value versus turbidity were generated, and

ordinary least-squares regression was applied. Extracted metrics included slope, intercept, coefficient of determination (R^2), p-value, standard deviation of residuals and number of collocated pairs (N). Comparative stability of these metrics across scenarios informed the robustness assessment [24].

Table 2 Representative parameter settings retained for detailed analysis. The Baseline scenario reflects physically conservative processing choices for shallow, nearshore waters. Subsequent scenarios isolate the effects of spatial buffering, temporal matching, and water-only masking.

Scenario	Buffer (m)	Time Window	Cloud Threshold	Water-only Mask
1. Baseline	200	±6 h	10%	ON
2. Spatial Sensitivity	300	±6 h	10%	ON
3. Temporal Sensitivity	200	±120 h	10%	ON
4. Masking Sensitivity	200	±6 h	10%	OFF

All processing and analysis were implemented in Python (script version 2), including automatic ingestion of CTD and Sentinel-2 data, depth filtering, Sentinel-Hub queries, reproducible masking/aggregation logic, logging of skipped records and checkpointing for safe resumption. Detailed logs, results tables and code are archived for full reproducibility [25]. The overall workflow is summarized in Figure 3.

3. Results

3.1. Depth-aggregated in-situ turbidity observations

To investigate vertical coherence of optical water properties, the CTD turbidity data were aggregated at three discrete near-surface depth layers (0–1 m, 1–2 m, and 2–3 m). For each sampling location, turbidity values were averaged over the respective layer and grouped by unique longitude/latitude coordinates, thereby reducing the influence of repeated casts or small-scale ship movement. The resulting datasets contained 53 independent stations per depth layer, each associated with a representative sampling time and location suitable for Sentinel–2 collocation.

A comparison of the three layers indicated that turbidity varied smoothly with depth and exhibited no abrupt gradients within the upper 3 m. This vertical coherence suggests that surface-reflectance measurements from Sentinel–2 should be sensitive to turbidity across the same range, thereby justifying the use of layer-wise collocations.

3.2. Relationships between Sentinel–2 indices and turbidity

Three broadband satellite indices were evaluated independently for each depth-aggregated turbidity dataset: the Normalized Difference Turbidity Index (NDTI), the Normalized Difference Water Index (NDWI), and the Normalized Difference Vegetation Index (NDVI). For all runs, Sentinel–2 L2A scenes were selected using a ±48-hour temporal window, a maximum scene cloud fraction of 40 %, and spatial buffer sizes of 120–400 m around each CTD station (Section 2). Collocations were performed at 20 m and 40 m pixel resolutions, with results being highly consistent across the entire parameter range (Figure 4; Table 1).

Across all depth layers, NDWI showed the clearest and most stable relationship with in-situ turbidity. At 1–2 m depth, NDWI explained approximately 13 % of the observed turbidity variability ($R^2 = 0.134$; $n = 53$), with a positive slope of 34.3 FTU per index unit. A very similar relationship was obtained for the 2–3 m layer ($R^2 = 0.141$; slope = 34.8 FTU unit^{-1}). These nearly identical slopes across adjacent depth layers indicate that NDWI is responding to an optically coherent water column signal rather than depth-specific noise.

NDVI exhibited a moderately strong but consistently negative relationship with turbidity, reflecting the well-understood attenuation of red–NIR contrast under elevated sediment loads. For the 1–2 m and 2–3 m layers, R^2 values ranged from 0.114 to 0.121, with regression slopes between −65 and −67 FTU unit^{-1}.

By contrast, NDTI demonstrated only weak correlations with turbidity in all configurations. Across both depth layers, R^2 remained below 0.01 (Table 1). This behavior suggests that the SWIR-based NDTI signal, which is theoretically linked to suspended sediment concentration, carries limited sensitivity to the relatively moderate turbidity range observed in this study area. In other words, NDTI does not appear to be a practical turbidity indicator for these coastal waters.

The resulting relationships between turbidity (FTU) and the three indices for the 1–2 m and 2–3 m layers are shown in Figure 4.

3.3. Influence of spatial buffer size and pixel resolution

To test robustness against parameter selection, all analyses were repeated using spatial buffers of 120, 200, 300, and 400 m and pixel resolutions of 20 and 40 m. Across these configurations, the resulting regression slopes and R^2 values varied by less than ±0.01 for all indices and depth layers. This minimal sensitivity indicates that the detected index–turbidity relationships are not artefacts of specific buffer choices or processing resolutions.

These values confirm that submerged vegetation or benthic contributions are suppressed at turbid stations, leading to systematically lower NDVI values.

The stability of NDWI and NDVI slopes across parameter settings further suggests that both indices respond primarily to inherent turbidity-driven changes in water-leaving reflectance, not to small-scale spatial heterogeneity or pixel-aggregation effects. The sensitivity of the index–turbidity regressions to spatial buffer size and pixel resolution is summarized in Figure 5.

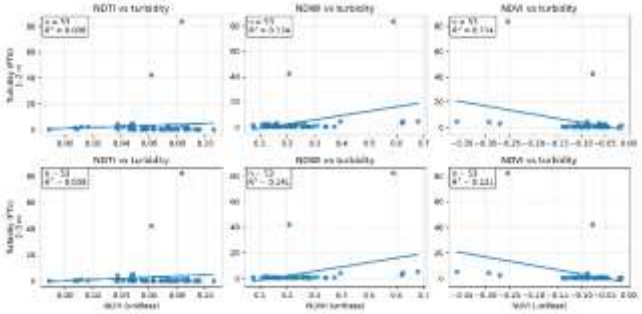

Figure 4 . Scatterplots of in-situ turbidity (FTU) versus three Sentinel–2 indices (NDTI, NDWI, NDVI) for two depth-aggregated layers (1–2 m and 2–3 m). Each point represents a spatially averaged CTD station (n = 53 per layer), collocated with the closest Sentinel–2 L2A scene within a ±48 h window. Regression lines (solid) are shown where n ≥ 2, together with the corresponding coefficients of determination R^2. NDWI exhibits the strongest and most internally consistent positive relationship with turbidity ($R^2 \approx 0.13$–0.14), while NDVI shows a systematic negative relationship ($R^2 \approx 0.11$– 0.12) reflecting the attenuation of red–NIR contrast under elevated turbidity. NDTI displays negligible sensitivity to the observed turbidity range ($R^2 < 0.01$), indicating that SWIR-based ratios are not informative for these coastal waters.

3.4. Summary of depth-based findings

Taken together, the depth-aggregated analysis demonstrates that:

1. NDWI provides the most consistent and interpretable proxy for turbidity (FTU) in these shallow coastal waters, with stable slopes and the highest R^2 values across depth layers.
2. NDVI provides a complementary signal, capturing the suppression of red–NIR contrast in more turbid conditions, relevant especially for mixed bottom–water pixels.
3. NDTI shows negligible sensitivity to the observed turbidity range and is not recommended for further development in this study area.
4. The vertical structure of turbidity (0–3 m) is sufficiently coherent that Sentinel–2 reflectance at the surface effectively integrates turbidity conditions across the upper water column.

This multi-layer consistency provides strong support for using NDWI-based or NDVI-based estimators as inputs for a semi-empirical turbidity retrieval model or for mapping relative turbidity gradients in the larger study region.

4. Discussion

4.1. Performance of Sentinel–2 indices for turbidity retrieval

The present analysis provides a detailed evaluation of three broadband Sentinel–2 indices (NDWI, NDVI, NDTI) for describing near-surface turbidity in a shallow coastal environment. Among the tested indices, NDWI exhibited the strongest and most consistent relationship with in-situ turbidity. This behavior reflects the physical sensitivity of the green and near-infrared bands to suspended particulate matter (SPM): increased turbidity enhances multiple scattering in the green wavelengths while simultaneously reducing NIR reflectance, producing a monotonic positive response in NDWI. The regression slopes for NDWI were remarkably stable across the 1–2 m and 2–3 m depth layers, suggesting that the index predominantly responds to depth-integrated optical conditions rather than depth-specific artefacts.

NDVI also showed a coherent turbidity signal but with systematically negative slopes and moderately lower explanatory power. This behavior arises because turbid waters attenuate red–NIR contrast, suppressing contributions from submerged vegetation or benthic reflectance. As a consequence, NDVI becomes useful for distinguishing clearer sites—where benthic signatures or submerged aquatic vegetation may still contribute to the reflectance—from high-turbidity stations where such signals are overwhelmed. In contrast, NDTI displayed negligible sensitivity to turbidity across all tested configurations. Although SWIR reflectance can theoretically respond to SPM, the combination of lower signal-to-noise ratios over water, atmospheric correction uncertainties, and the moderate turbidity range observed in this study likely prevented NDTI from exhibiting meaningful variation.

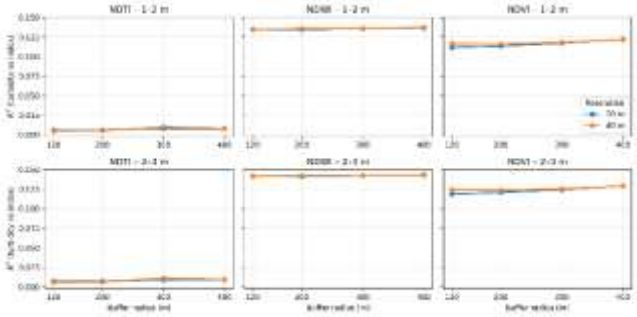

Figure 5 Sensitivity of the index–turbidity relationships to spatial buffer size and pixel resolution. Panels show the coefficient of determination R^2 for linear regressions between turbidity (FTU) and each Sentinel–2 index (NDTI, NDWI, NDVI) as a function of buffer radius (120–400 m), for two depth layers (1–2 m and 2–3 m; rows) and two processing resolutions (20 m and 40 m; separate curves). For all indices and depths, R^2 varies by less than approximately 0.01 across the tested parameter range, indicating that the identified NDWI- and NDVI-based turbidity signals are robust to moderate changes in buffer size and pixel resolution.

4.2. Influence of depth structure and aggregation

A key result of this study is the vertical coherence of turbidity in the upper 3 m of the water column, as demonstrated by averaged CTD profiles and depth-binned statistics. The similarity of index–turbidity relationships across the 1–2 m and 2–3 m layers suggests that surface reflectance measured by Sentinel–

2 is representative of optical conditions throughout the mixed near-surface layer. This coherence justifies the use of layer-averaged turbidity values for satellite collocation and indicates that the spatial and temporal variability observed at the surface is largely driven by depth-integrated processes such as resuspension, sediment transport, and optical scattering by fine particulate matter.

Depth aggregation also reduced within-station variability, enabling a cleaner comparison between in-situ measurements and pixel-scale satellite retrievals. Because the vessel revisited several stations multiple times, the aggregation procedure ensured that each geographic location contributed a single representative turbidity value to the collocation dataset while preserving the natural spatial gradients present in the domain.

4.3. Sensitivity to spatial buffer size and pixel resolution

The robustness of the index–turbidity relationships across buffer radii and processing resolutions is an important outcome for operational monitoring. Tests using four spatial buffers (120–400 m) and two-pixel resolutions (20 m and 40 m) showed that the resulting R^2 values varied by less than approximately 0.01 across all configurations. This minimal sensitivity indicates that the turbulence, mixing, and sediment-transport regimes in the study area produce spatially smooth turbidity gradients at scales larger than a few hundred meters. Consequently, satellite-derived turbidity products need not rely on a narrowly optimized buffer size; moderate variation in spatial support produces consistent results.

This behavior contrasts with environments characterized by sharp turbidity fronts, strong plume gradients, or complex bathymetry, where small changes in buffer size can significantly affect representativeness. The present findings suggest that the optical environment of the study area is sufficiently homogeneous at the sub-kilometer scale to support robust Sentinel–2-based monitoring.

4.4. Implications for semi-empirical turbidity retrieval

The depth-robust and spatially stable response of NDWI to turbidity suggests that the index can serve as a reliable predictor for semi-empirical turbidity retrieval models. While the explained variance is moderate ($R^2 \approx 0.13$–0.14), the stability of the regression slopes across depth layers and buffer sizes indicates that the physical relationship is consistent and potentially translatable to other seasons or adjacent regions with similar optical conditions. Future work could incorporate non-linear or multi-index regression techniques to capture additional spectral sensitivity to SPM, or integrate physics-informed corrections such as depth-dependent light attenuation or benthic reflectance suppression.

NDTI's uniformly weak performance highlights that SWIR-based methods may be unsuitable for moderate turbidity regimes in shallow coastal water. Instead, methods based on green/NIR contrasts, like NDWI, appear to be better suited for environments where SPM concentrations are not high enough for SWIR reflectance to rise above noise levels.

NDTI's uniformly weak performance highlights that SWIR-based methods may be unsuitable for moderate turbidity regimes in shallow coastal water. Instead, methods based on green/NIR contrasts, like NDWI, appear to be better suited for environments where SPM concentrations are not high enough for SWIR reflectance to rise above noise levels.

4.5. Limitations and opportunities

Several limitations merit consideration. First, the temporal matching window of ±48 h, although necessary to ensure adequate Sentinel–2 coverage, introduces the possibility of decorrelation under rapidly varying conditions. Nonetheless, the observed stability of the index–turbidity relationships suggests that the temporal variability in turbidity during the sampling period was not sufficient to obscure the satellite signal. Second, the CTD turbidity values were measured in FTU, and although this unit is functionally comparable to NTU for field applications, the exact equivalence may vary across sensors and calibration standards. Third, the study was conducted during a single observational campaign; future research using multi-season or multi-year datasets would help test the temporal generality of the findings.

Despite these limitations, the results demonstrate that Sentinel–2 provides a consistent and reproducible turbidity signature in shallow, moderately turbid environments. The combination of depth-aggregated in-situ profiles, buffer-insensitivity, and stable index performance underscores the potential for using Sentinel–2 for routine turbidity monitoring in similar coastal regions.

5. Conclusions

This study evaluated the ability of three Sentinel–2 spectral indices—NDWI, NDVI, and NDTI—to represent near-surface turbidity in a shallow coastal environment using a depth-aggregated CTD dataset. By averaging turbidity measurements over the upper 1–3 m and collocating them with spatially buffered Sentinel–2 scenes, I quantified the sensitivity of each index to suspended particulate matter under realistic oceanographic conditions.

Among the tested indices, NDWI consistently showed the strongest and most stable relationship with turbidity across all depth layers, spatial buffers, and processing resolutions. NDVI also reflected turbidity variations, but with weaker explanatory power and a negative

response linked to red–NIR attenuation in turbid waters. NDTI, based on SWIR wavelengths, demonstrated negligible predictive value for the turbidity range encountered in this environment.

The index–turbidity relationships were remarkably robust to changes in buffer radius (120–400 m) and pixel resolution (20 m vs. 40 m), indicating that the optical environment is spatially smooth at sub-kilometer scales. This result supports the feasibility of operational turbidity monitoring without strict parameter optimization. The vertical coherence of turbidity within the upper 3 m further supports the use of depth-averaged in-situ values for satellite validation. Overall, the findings highlight NDWI as a reliable semi-empirical proxy for turbidity in shallow, moderately turbid coastal waters, while underscoring limitations of SWIR-based indices such as NDTI. The methodological framework developed here—including depth aggregation, temporal collocation, and robust sensitivity analysis—provides a transferable basis for turbidity mapping in similar environments.

Future work should explore multi-season datasets, incorporate physics-based reflectance corrections, and evaluate synergistic approaches combining Sentinel–2 with higher-frequency missions (e.g. Sentinel–3 and PlanetScope) to improve temporal coverage and predictive strength. Together, these advancements may enhance regional water-quality monitoring and contribute to sustainable coastal management.

Overall, the limited sensitivity of turbidity retrieval performance to commonly used satellite–in-situ collocation parameters demonstrates that robust and reproducible Sentinel–2–based estimates can be achieved under realistic operational constraints, supporting routine and operational coastal water-quality monitoring.

Acknowledgment

The author gratefully acknowledges the support of the Iranian National Institute for Oceanography and Atmospheric Sciences, for facilitating the execution of the Persian Gulf Explorer research cruises across the Persian Gulf and Gulf of Oman.

6. References

1. Kirk, J.T.O. Light and Photosynthesis in Aquatic Ecosystems, 3 ed.; Cambridge University Press: Cambridge, UK, 2011.
2. Dennison, W.C.; Orth, R.J.; Moore, K.A.; Stevenson, J.C.; Carter, V.; Kollar, S.; Bergstrom, P.W.; Batiuk, R.A. Assessing Water Quality with Submersed Aquatic Vegetation: Habitat Requirements as Barometers of Chesapeake Bay Health. BioScience 1993,43, 86–94. https://doi.org/10.2307/1311969.
3. Cloern, J.E. Our Evolving Conceptual Model of the Coastal Eutrophication Problem. Marine Ecology Progress Series 2001, 210, 223–253. https://doi.org/10.3354/meps210223.
4. Morel, A.; Prieur, L. Analysis of Variations in Ocean Color. Limnology and Oceanography 1977, 22, 709–722. https://doi.org/10.4319/lo.1977.22.4.0709.
5. Binding, C.E.; Bowers, D.G.; Mitchelson-Jacob, E.G. Estimating Suspended Sediment Concentrations from Ocean Colour Measurements in Moderately Turbid Waters: The Impact of Variable Particle Scattering Properties. Remote Sensing of Environment 2005, 94, 373–383. https://doi.org/10.1016/j.rse.2004.11.002.
6. Nechad, B.; Ruddick, K.G.; Neukermans, G. Calibration and Validation of a Generic Multisensor Algorithm for Mapping of Total Suspended Matter in Turbid Waters. Remote Sensing of Environment 2010, 114, 854–866. https://doi.org/10.1016/j.rse.2009.11.022.
7. Drusch, M.; Del Bello, U.; Carlier, S.; Colin, O.; Fernandez, V.; Gascon, F.; Hoersch, B.; Isola, C.; Laberinti, P.; Martimort, P.; et al. Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. Remote Sensing of Environment 2012, 120, 25–36. https://doi.org/10.1016/j.rse.2011.11.026.
8. Pahlevan, N.; Sarkar, S.; Franz, B.A.; Balasubramanian, S.V.; He, M. Sentinel-2 MultiSpectral Instrument (MSI) Data Processing for Aquatic Science Applications: Demonstrations and Validations. Remote Sensing of Environment 2017, 201, 47–56. https://doi.org/10.1016/j.rse.2017.08.033.
9. Novoa, S.; Doxaran, D.; Ody, A.; Vanhellemont, Q.; Lafon, V.; Lubac, B.; Gernez, P. Atmospheric Corrections and Multi-Conditional Algorithm for Multi-Sensor Remote Sensing of Suspended Particulate Matter in Low-to High-Turbidity Coastal Waters. Remote Sensing 2017, 9, 61. https://doi.org/10.3390/rs9010061.
10. Dogliotti, A.I.; Ruddick, K.; Nechad, B.; Doxaran, D.; Knaeps, E. A Single Algorithm to Retrieve Turbidity from Remotely-Sensed Data in All Coastal and Estuarine Waters. Remote Sensing of Environment 2015, 156, 157–168. https://doi.org/10.1016/j.rse.2014.09.020.
11. Brando, V.E.; Anstee, J.M.; Wettle, M.; Dekker, A.G.; Phinn, S.R.; Roelfsema, C. A Physics Based Retrieval and Quality Assessment of Bathymetry from Suboptimal Hyperspectral Data. Remote Sensing of Environment 2009, 113,

755–770. https://doi.org/10.1016/j.rse.2008.12.003.
12. Lee, Z.; Carder, K.L.; Mobley, C.D.; Steward, R.G.; Patch, J.S. Hyperspectral Remote Sensing for Shallow Waters: I. A Semianalytical Model. Applied Optics 1998, 37, 6329–6338. https://doi.org/10.1364/AO.37.006329.
13. Hedley, J.D.; Harborne, A.R.; Mumby, P.J. Technical Note: Simple and Robust Removal of Sun Glint for Mapping Shallow-Water Benthos. International Journal of Remote Sensing 2005, 26, 2107–2112. https://doi.org/10.1080/01431160500034086.
14. Moses, W.J.; Gitelson, A.A.; Berdnikov, S.; Povazhnyy, V. Operational MERIS-Based NIR-Red Algorithms for Estimating Chlorophyll-a Concentrations in Coastal Waters – The Azov Sea Case Study. Remote Sensing of Environment 2012, 121, 118–124. https://doi.org/10.1016/j.rse.2012.01.024.
15. Werdell, P.J.; Bailey, S.W. An Improved In-Situ Bio-Optical Data Set for Ocean Color Algorithm Development and Satellite Data Product Validation. Remote Sensing of Environment 2005, 98, 122–140. https://doi.org/10.1016/j.rse.2005.07.001.
16. McFeeters, S.K. The Use of the Normalized Difference Water Index (NDWI) in the Delineation of Open Water Features. International Journal of Remote Sensing 1996, 17, 1425–1432. https://doi.org/10.1080/01431169608948714.
17. Lacaux, J.P.; Tourre, Y.M.; Vignolles, C.; Ndione, J.A.; Lafaye, M. Classification of Ponds from High-Spatial Resolution Remote Sensing: Application to Rift Valley Fever Epidemics in Senegal. Remote Sensing of Environment 2007, 106, 66–74. https://doi.org/10.1016/j.rse.2006.07.012.
18. Brezonik, P.L.; Olmanson, L.G.; Finlay, J.C.; Bauer, M.E. Factors Affecting the Measurement of CDOM by Remote Sensing of Optically Complex Inland Waters. Remote Sensing of Environment 2015, 157, 199–215. https://doi.org/10.1016/j.rse.2014.04.033.
19. Lee, Z.; Carder, K.L.; Arnone, R.; et al. From deep open ocean to coastal waters: A seamless whole-water optical property model. Remote Sensing of Environment 2019, 235, 111–196. https://doi.org/10.1016/j.rse.2019.111196.
20. Waycott, M.; Duarte, C.M.; Carruthers, T.J.B.; Orth, R.J.; Dennison, W.C.; Olyarnik, S.; Calladine, A.; Fourqurean, J.W.; Heck, K.L.J.; Hughes, A.R.; et al. Accelerating loss of seagrasses across the globe threatens coastal ecosystems. Proceedings of the National Academy of Sciences of the United States of America 2009, 106, 12377–12381. https://doi.org/10.1073/pnas.0905620106.
21. Lu, J.; Li, X.; Li, Y.; Zhang, Y. A semi-analytical retrieval algorithm of water quality parameters in shallow waters with bottom effects. Remote Sensing 2015, 7, 995–1015. https://doi.org/10.3390/rs70100995.
22. Zhang, Y.; Shi, K.; Zhou, Y.; Song, K.; Liu, X. Spatial–temporal variability and driving forces of satellite-derived lake turbidity in large inland lakes. Remote Sensing of Environment 2018, 209, 395–407. https://doi.org/10.1016/j.rse.2018.02.062.
23. Kutser, T. The importance of selecting the correct sensitivity analysis method for optical remote sensing of water quality. International Journal of Remote Sensing 2012, 33, 5569–5576. https://doi.org/10.1080/01431161.2012.666812.
24. Willmott, C.J.; Matsuura, K. Advantages of the Mean Absolute Error (MAE) over the Root Mean Square Error (RMSE) in Assessing Average Model Performance. Climate Research 2005, 30, 79–82. https://doi.org/10.3354/cr030079.
25. Hollender, J.; van Bavel, B.; Dulio, V.; Slobodnik, J.; Munthe, J. Toward a review of reproducible environmental science: Best practices and challenges. Environmental Sciences Europe 2019, 31, 51. https://doi.org/10.1186/s12302-019-0230-6.