Marine Pollution Bulletin 231 (2026) 120004

Contents lists available at ScienceDirect

Marine Pollution Bulletin

journal homepage: www.elsevier.com/locate/marpolbul

Monitoring network design limits detection of sediment-mediated contaminant mobilisation in the River Mersey estuary (UK)

Deelaram Nangir[a,*], Manolia Andredaki[a], Iacopo Carnacina[a]

[a] School of Civil Engineering and Built Environment, Liverpool John Moores University, Liverpool, United Kingdom

ARTICLE INFO

Keywords:
Estuarine pollution
Sediment contaminants
Monitoring network
Machine learning
Water quality monitoring
River Mersey

ABSTRACT

Suspended sediments regulate turbidity and act as major vectors for metals, organic contaminants, and nutrient fluxes in estuarine systems. Effective assessment of sediment-mediated pollution therefore depends on coordinated monitoring of sediment and contaminant dynamics. This study evaluates whether a long-term regulatory monitoring network can resolve sediment-associated pollution dynamics in the River Mersey estuary (UK).

A total of 44,516 water-column observations (2015–2025) across 42 monitoring stations and 187 determinants were analysed. Monitoring frequency classification revealed pronounced structural imbalance: only 6.9% of determinants were sampled routinely ($\geq$1000 observations), whereas 67.4% were recorded fewer than 100 times. Physical and oxygen-related variables accounted for 52% of all observations, while sediment and turbidity indicators comprised only 4.2%.

Temporal separation analysis revealed a bimodal sampling structure characterised by episodic co-sampling embedded within extended periods of asynchronous measurement. While pairwise co-sampling coherence (CCR) was often moderate to high across several process groups, strict multivariate coherence (CCR_{strict}) remained low across much of the network, indicating that complete multivariate observation states suitable for integrated sediment–contaminant assessment were uncommon.

Diagnostic machine-learning reconstruction yielded moderate performance under random validation ($R^2 =$ 0.588) but substantially lower spatial generalisation under grouped cross-validation (mean $R^2 \approx 0.26$), indicating limited transferability across stations. These findings suggest that sediment-mediated pollution assessment is constrained more strongly by monitoring architecture, temporal alignment, and environmental heterogeneity than by modelling approaches alone. The study introduces a diagnostic framework combining temporal separation analysis, coherence metrics, and hierarchical validation for evaluating monitoring-network suitability in estuarine systems.

1. Introduction

Estuaries act as long-term sinks for industrial contaminants including heavy metals, hydrocarbons and legacy organic pollutants, many of which remain stored in sediments and can be remobilised during resuspension events. Effective pollution surveillance therefore depends not only on chemical monitoring but also on the capacity of monitoring networks to capture sediment-mediated contaminant mobilisation (Defeo and Elliott, 2021; Robins et al., 2016). They are subject to strong physical forcing, complex biogeochemical interactions, and sustained human pressures, including wastewater discharges, industrial activity, navigation, and dredging (Greenwood et al., 2019; Vörösmarty et al., 2010). As focal points for water-quality management and regulatory compliance, estuaries require monitoring systems capable of resolving the processes that govern pollutant transport and ecological response. Among these processes, sediment dynamics play a central role by controlling light attenuation and mediating the transport, storage, and remobilisation of metals, organic contaminants, and nutrients (Grandjean et al., 2024; Lunt and Smee, 2020).

Because suspended sediments bind and remobilise contaminants, the capacity to resolve sediment variability directly constrains the detection of pollutant transport pathways in industrialised estuaries (Eggleton and Thomas, 2004; Turner and Millward, 2002). In dynamic tidal systems, contaminant mobilisation frequently occurs during short-lived resuspension events, during which contaminants may rapidly exchange between particulate and dissolved phases through processes including

* Corresponding author.
E-mail address: d.nangir@ljmu.ac.uk (D. Nangir).

https://doi.org/10.1016/j.marpolbul.2026.120004
Received 9 March 2026; Received in revised form 4 June 2026; Accepted 9 June 2026
Available online 17 June 2026

adsorption-desorption dynamics, salinity-driven partitioning, and flocculation (Eggleton and Thomas, 2004; Turner and Millward, 2002). These processes can operate over tidal to sub-daily timescales, meaning that asynchronous or infrequent sediment sampling may systematically fail to capture critical pollution episodes and transient contaminant mobilisation states (Geyer and MacCready, 2014). Despite this importance, sediment processes remain among the most difficult components of estuarine systems to characterise using long-term monitoring data (Matos et al., 2024).

Regulatory water-quality monitoring networks rely on in-situ measurements collected at fixed stations over extended periods. These networks provide essential long-term records of chemical and physical parameters and underpin management and policy decisions (Lonsdale et al., 2022). However, sediment-related indicators such as turbidity and suspended particulate matter are often sampled less frequently, at fewer stations, and under different monitoring objectives than routinely measured variables including temperature, dissolved oxygen, or biochemical oxygen demand (Karthick et al., 2024; Skarbøvik et al., 2023). This imbalance may create systematic blind spots in the detection of sediment-associated contaminant pulses, particularly during episodic high-flow or resuspension events. Consequently, sediment-bound contaminant dynamics are systematically under-represented in long-term regulatory datasets, even in estuaries where sediment transport dominates water-quality variability (Matos et al., 2024).

In parallel with the expansion of monitoring archives, advances in machine learning and data-driven modelling have stimulated efforts to reconstruct unmeasured or sparsely measured water-quality parameters from co-observed variables. Under coordinated sampling conditions, machine-learning models have successfully reconstructed sediment indicators from multivariate physical and chemical datasets, including frameworks integrating satellite remote sensing under aligned temporal acquisition (Chowdhury et al., 2023; Kumar et al., 2022; Liu et al., 2025; Magrì et al., 2023; Nangir et al., 2025). However, these demonstrations are typically based on datasets explicitly designed to ensure multivariate temporal coherence rather than compliance-oriented regulatory archives.

Most machine-learning studies implicitly assume that in-situ datasets contain temporally coherent multivariate observations, meaning that multiple parameters are measured concurrently or within sufficiently narrow time windows to represent a single environmental state (Akhlaq et al., 2024; Zhu et al., 2022a). In regulatory programmes structured around discrete compliance objectives, this assumption is rarely tested. Structural misalignment between sediment and contaminant measurements may therefore remain undetected within routine surveillance frameworks. When predictive performance is poor, limitations are commonly attributed to algorithm selection, data volume, or measurement noise rather than to monitoring architecture itself (Chicco et al., 2021). Yet if sediment sampling is temporally decoupled from contaminant measurements, even sophisticated modelling approaches cannot compensate for the absence of co-observed pollutant transport dynamics.

Monitoring activities in highly dynamic estuarine environments are frequently distributed across multiple programmes with distinct regulatory or investigative objectives. Sediment sampling may be conducted episodically for compliance checks or targeted studies, while routine water-quality monitoring focuses on different parameter subsets and follows independent temporal schedules (Greenwood et al., 2019; Lonsdale et al., 2022). Such temporal decoupling reduces the probability that sediment observations coincide with peak contaminant mobilisation, thereby limiting the ability of monitoring networks to capture pollution-relevant state transitions. Despite its direct implications for marine pollution surveillance and compliance assessment, the capacity of regulatory monitoring networks to support sediment-mediated contaminant detection has received limited systematic evaluation.

The River Mersey estuary provides a robust test case for examining these issues. As a heavily urbanised and industrialised estuary, the Mersey has a documented legacy of metal and organic contamination and remains subject to active pollution management (Burton, 2003; Hawkins et al., 2020; Langston et al., 2006). During the nineteenth and twentieth centuries, extensive industrial discharge resulted in substantial accumulation of heavy metals and organic contaminants within estuarine sediments. Although water quality has improved following regulatory interventions, contaminated sediments remain an important secondary pollution source. In this macro-tidal system, tidal resuspension events can remobilise legacy contaminants into the water column, making sediment–water coupling central to pollution surveillance and environmental risk assessment. It has been monitored extensively under regulatory frameworks for multiple decades and represents a data-rich estuarine system by international standards. Whether its long-term in-situ monitoring archive is adequately configured to detect sediment-mediated contaminant transport at the system scale remains unresolved.

This study reframes sediment reconstruction as a pollution-detection capacity problem rather than a modelling optimisation exercise. We hypothesise that temporal misalignment between sediment and contaminant monitoring imposes structural limits on pollution detectability that cannot be resolved through increased modelling complexity alone. Using a decade of regulatory in-situ data from the River Mersey (2015–2025), we systematically quantify the distribution, spatial coverage, and temporal alignment of sediment-related observations relative to routinely monitored chemical and physical parameters. Machine learning is applied as a structural diagnostic of monitoring configuration rather than as a predictive optimisation tool, in order to evaluate whether existing monitoring architectures contain sufficient multivariate structure to support transferable sediment reconstruction across stations.

Accordingly, this study evaluates whether long-term regulatory monitoring networks are structurally configured to resolve sediment-mediated contaminant transport at the system scale. This is achieved by (i) quantifying the frequency distribution and spatial coverage of sediment-related determinants relative to routinely monitored parameters, (ii) assessing temporal co-occurrence and multivariate coherence across stations, and (iii) employing diagnostic machine-learning frameworks to test reconstruction feasibility and cross-station transferability under existing monitoring configurations. By linking sampling architecture directly to pollutant detection capacity, the study establishes a transferable framework for evaluating whether regulatory monitoring investments are sufficient to resolve sediment-associated contaminant risk in dynamic estuarine systems.

2. Materials and methods

2.1. Data source and accessibility

All in-situ water-quality data analysed in this study were obtained from the UK Environment Agency's open-access Water Quality Archive (Environment Agency, 2025). Data were downloaded on 28 December 2025 to ensure a fixed and reproducible archive state.

This repository provides quality-controlled measurements collected under regulatory, operational, and investigative monitoring programmes across England, covering rivers, estuaries, coastal waters and associated discharge points. The archive includes a wide range of chemical, physical and biological determinants, together with detailed station-level metadata describing sampling location, sample material type, analytical method and sampling purpose (Lonsdale et al., 2022).

Data was accessed through the Environment Agency Water Quality Explorer platform. Records were filtered to include only water-column samples collected at monitoring stations located within the River Mersey estuary and its immediately adjacent coastal reaches. The estuarine boundary was defined using station metadata descriptors and geographic coordinates corresponding to the tidal Mersey between Warrington and Liverpool Bay, consistent with regulatory station classification within the archive. Samples associated with non-aquatic media (e.g. sediments, biota, or effluents) or unrelated monitoring objectives were excluded. Only observations with valid numeric results and clearly defined sampling timestamps (phenomenon Time) were retained.

Determinants were selected without pre-screening by chemical class in order to preserve the full structure of the regulatory monitoring archive. This inclusive selection enables evaluation of whether sediment observations temporally coincide with contaminant measurements within the existing surveillance configuration, rather than optimising parameter subsets for modelling performance.

Duplicate records were removed during data preparation. Duplicates were strictly defined as entries sharing identical combinations of station identifier, determinant name, sampling timestamp, and reported numeric value. This conservative definition ensured that legitimate replication or confirmatory measurements were not inadvertently removed.

No temporal aggregation, interpolation, smoothing, or gap-filling was applied. All analyses were conducted on the archive as collected, ensuring that subsequent coherence and reconstruction diagnostics reflect monitoring design rather than post-processed regularisation.

The use of an openly accessible regulatory archive, combined with explicit filtering criteria, ensures transparency and full reproducibility. No proprietary datasets were used.

2.2. Study area and station selection

The River Mersey estuary was selected as the focus of this study due to its long history of anthropogenic pressure, extensive regulatory monitoring record and strong physical and biogeochemical variability (Burton, 2003; Hawkins et al., 2020; Hurley et al., 2017). Historically impacted by industrial discharges and urbanisation, the Mersey has undergone substantial environmental recovery, while remaining one of the most intensively managed estuaries in the United Kingdom in terms of wastewater discharge regulation, navigation, dredging and emerging renewable energy infrastructure (Becker et al., 2017; Greenwood et al., 2019; Hawkins et al., 2020). The estuary is macro-tidal, highly turbid, and characterised by strong tidal resuspension processes that directly influence sediment-associated contaminant transport, making it a suitable system for evaluating whether existing monitoring design adequately captures sediment-mediated pollutant mobilisation under dynamic forcing.

Monitoring stations located within the estuarine and immediately adjacent coastal reaches of the River Mersey were identified using spatial metadata and station descriptors provided in the archive. The estuarine footprint was defined based on tidal influence and regulatory station classification, extending from the tidal limit near Warrington to the mouth of the estuary at Liverpool Bay, including nearshore coastal stations hydrodynamically connected to the estuarine plume. This spatial identification step yielded 97 regulatory monitoring stations within the defined footprint (Fig. 1).

Only stations associated with water-column sampling (sample material classified as "water" or "unfiltered water") were retained. Stations dedicated exclusively to non-aquatic media (e.g., sediment cores, biota, soils) or to non-water monitoring objectives (e.g., atmospheric deposition, solid waste, or terrestrial investigations) were excluded. Stations lacking consistent spatial metadata or valid water-quality observations during the study period were also removed.

This filtering resulted in a final dataset comprising 42 regulatory in-situ monitoring stations distributed across the estuarine system. The reduction from 97 to 42 stations reflects exclusion of non-water sampling locations and sites without continuous water-column measurements during the study period, rather than selective removal based on sediment availability.

This diversity reflects the operational structure of the regulatory monitoring network and is essential for evaluating how monitoring objectives influence sediment data availability (Environment Agency, 2025; Lonsdale et al., 2022).

Fig. 1. Study area and spatial distribution of regulatory in-situ monitoring stations within the River Mersey estuary. Monitoring stations ($n = 97$) are shown overlaid on high-resolution satellite imagery. The inset map provides regional context and highlights the location of the study area within the United Kingdom.

The retained stations encompassed routine surveillance sites, compliance-focused discharge monitoring locations, and episodic investigative stations. Inclusion of heterogeneous station functions was intentional, as the objective was to evaluate the structural coherence of the full regulatory monitoring system rather than an optimised research subset. Excluding compliance or investigative stations would have artificially increased apparent structural coherence and therefore masked the operational realities of regulatory monitoring design. Full station metadata are provided in Supplementary Material S1, including station identifier (Station ID), station name, and total number of observations per station (n_samples).

2.3. Temporal scope and data preparation

The analysis covered a ten-year period from January 2015 to January 2025. All observations were standardised to a common timestamp format using the archive field "phenomenonTime".

Duplicate records were defined as entries sharing the same monitoring station identifier, determinant name, sampling timestamp, and reported result value. Where such exact duplicates occurred, a single instance was retained.

Non-numeric result fields (e.g., qualitative descriptors, censored values, presence/absence indicators) were excluded at the determinant level. Censored values reported below analytical detection limits were also excluded where no numeric surrogate value was provided in the archive. Stations or sampling events were not removed solely on the basis of partial non-numeric entries, ensuring preservation of the structural sampling record.

No gap-filling, temporal interpolation, or aggregation was applied. This decision ensures that any detected temporal fragmentation or multivariate incompleteness reflects inherent monitoring design rather than artefacts of data processing. (Bieroza et al., 2018; Zhu et al., 2022b).

All timestamps were retained at their original resolution, and no resampling to fixed temporal intervals was performed. Analyses therefore reflect the irregular sampling structure characteristic of regulatory monitoring programmes.

2.4. Parameter classification and grouping

To support system-level analysis and quantify monitoring priorities, determinants were classified using a dual framework based on monitoring frequency and process relevance, following approaches commonly used in large-scale water-quality assessments (Karthick et al., 2024; Kumar et al., 2022; Matos et al., 2024).

2.4.1. Monitoring frequency classification

Determinants were grouped into three frequency-based classes according to their total number of observations across the ten-year study period:

- Core (routine): $\geq$1000 observations
- Episodic: 100–999 observations
- One-off: $<$100 observations

Given the monitoring footprint (97 identified stations across ten years), the $\geq$1000 threshold corresponds approximately to $\geq$10 observations per year distributed across the network, indicating sustained system-level monitoring effort rather than isolated site-specific campaigns.

This threshold was selected to distinguish determinants embedded in sustained regulatory programmes from those measured intermittently or opportunistically. Sensitivity testing using moderate threshold variation ($\pm$20%) did not materially alter proportional class distributions, indicating robustness of the classification scheme.

These thresholds were designed to reflect structural monitoring commitment rather than impose rigid regulatory definitions.

2.4.2. Process-based grouping

Determinants were subsequently assigned to seven process-based groups reflecting their environmental role and relevance to estuarine water-quality assessment, consistent with established classifications in estuarine monitoring and sediment research (Hurley et al., 2017; Robins et al., 2016; Skarbøvik et al., 2023).

- Physical and hydromorphological conditions, including water temperature, pH, salinity, conductivity, water depth, sample depth, and tidal timing.
- Sediment and turbidity, including turbidity, suspended solids, visible oil or grease, and related particulate descriptors.
- Oxygen and organic pollution, including dissolved oxygen, biochemical oxygen demand (BOD), chemical oxygen demand (COD), and dissolved organic carbon.
- Nutrients and productivity, including nitrogen species, phosphorus compounds, silicate, chlorophylls, and phytoplankton-related indicators.
- Metals and inorganic contaminants, including dissolved and total metals, major ions, and alkalinity.
- Organic micropollutants, including polycyclic aromatic hydrocarbons, pesticides, industrial solvents, phenols, flame retardants, *per*- and polyfluoroalkyl substances, and other emerging contaminants.
- Biological indicators, including living biological components and pigment-based metrics providing ecological context.

Assignment was performed using determinant metadata descriptors and analytical method classifications provided in the archive. Classification rules were applied consistently across all determinants, and each determinant was assigned to a single primary process group to avoid double counting in monitoring effort quantification. For each process group, total observation count and number of unique determinants were calculated.

2.5. Station function and sampling intent

Station-level metadata describing sample material type and sampling purpose were analysed to characterise the intended function of each monitoring location. Classification was based on the Environment Agency "sampling purpose" and "monitoring programme" metadata fields, which describe the regulatory context under which each sample was collected. Stations were classified as primarily supporting routine surveillance when samples were collected under long-term, recurring monitoring programmes aimed at baseline water-quality assessment; as regulatory compliance stations when sampling was explicitly linked to discharge permits, consent conditions, or enforcement activities; and as episodic investigative stations when sampling was associated with short-term studies, incident response, or targeted investigations (Lonsdale et al., 2022). Where stations operated under multiple programme types, dominant classification was determined based on the majority of sampling records during the study period.

Sediment-capable stations were identified based on the presence of at least one sediment- or turbidity-related determinant during the study period. Stations lacking any sediment-related measurements were classified as non-sediment-capable (Defeo and Elliott, 2021). This distinction ensured that observed gaps in sediment monitoring were interpreted in the context of station design and monitoring intent, rather than as artefacts of data filtering (Matos et al., 2024).

Within the analytical water-column network (n = 42 stations), 28 stations recorded at least one sediment-related observation during the study period. To ensure sufficient temporal representation for diagnostic modelling, a threshold of $\geq$60 sediment observations per station was applied, corresponding approximately to an average of six sediment observations per year over ten years. Application of this criterion reduced the sediment-capable subset to 13 stations.

The ≥60-observation threshold was selected to balance the need for sufficient temporal coverage against the inherently sparse nature of sediment monitoring in regulatory networks. This threshold ensures that each retained station contains multiple seasonal cycles and interannual variability sufficient for cross-validation and grouped model evaluation. Smaller sample counts were found to produce unstable variance estimates and unreliable cross-station generalisation. Sensitivity testing using alternative cut-offs (≥40 and ≥ 80 observations) did not materially alter model transferability patterns, indicating that conclusions were not threshold-dependent.

2.6. Temporal coherence analysis

To assess the feasibility of integrated sediment analysis, temporal coherence between sediment indicators and other routinely monitored process groups was quantified. For each sediment observation, the minimum temporal separation Δt to the nearest observation from each non-sediment process group at the same station was calculated as:

$$\Delta t = \min(|t_s - t_i|) \tag{1}$$

where t_s is the timestamp of a sediment observation and t_i represents timestamps of observations from other process groups at the same station. Cross-station temporal matching was not performed.

To formalise monitoring alignment, we define temporal coherence as the proportion of sediment observations that possess at least one co-occurring measurement from another process group within a defined tolerance window. This proportion is termed the Co-sampling Coherence Ratio (CCR):

$$\text{CCR} = \text{N\_coherent}/\text{N\_total} \tag{2}$$

where N_coherent represents the number of sediment observations satisfying the temporal matching criterion and N_total represents the total number of sediment observations at the station.

CCR provides a dimensionless measure of pairwise monitoring coherence, quantifying the extent to which sediment observations align temporally with individual process groups. However, pairwise co-occurrence alone does not guarantee the existence of complete multivariate records suitable for integrated analysis or machine-learning reconstruction.

To evaluate analytical usability more directly, a stricter coherence metric ($\text{CCR}_{\text{strict}}$) was defined:

$$\text{CCR}_{\text{strict}} = \text{N\_complete}/\text{N\_total} \tag{3}$$

where Ncomplete represents the number of sediment observations for which all variables included within the final multivariate reconstruction framework were simultaneously available within the same sampling event.

Whereas CCR measures monitoring coherence between sediment observations and individual process groups, $\text{CCR}_{\text{strict}}$ measures the availability of complete multivariate observation states suitable for integrated reconstruction.

No universally established thresholds exist for CCR, as the metric is introduced here as a novel diagnostic index. The following indicative ranges are proposed based on the empirical relationship between CCR values and grouped cross-validation performance observed across the six modelling-eligible stations in this study, consistent with the principle that multivariate coherence should be evaluated relative to reconstruction objectives (Behmel et al., 2016; Roberts et al., 2017).

Roberts et al. (2017) demonstrated that the gap between random and spatially grouped cross-validation performance reflects the extent to which model skill depends on shared local structure rather than transferable system-wide relationships. CCR provides a monitoring-design metric that anticipates this gap: when CCR is low, co-sampling is insufficient to encode cross-station sediment dynamics, and the transferability gap is expected to be large. Accordingly, CCR values <0.2 are interpreted as critically low coherence and high risk of process underrepresentation; CCR values between 0.2 and 0.5 indicate moderate coherence where episodic co-sampling provides limited but potentially usable process information; and CCR values >0.5 indicate conditions under which multivariate reconstruction may become feasible at favourable stations.

These thresholds apply only to the pairwise CCR metric and are diagnostic rather than prescriptive. Cross-system validation using comparable regulatory archives is required before they can be considered transferable standards. As a more stringent criterion, $\text{CCR}_{\text{strict}}$ is expected to yield lower values than CCR due to the requirement for complete variable alignment.No equivalent thresholds are defined for $\text{CCR}_{\text{strict}}$, which is interpreted descriptively as a measure of complete multivariate state availability.

Three symmetric temporal matching windows were evaluated (same-day, ±3 days, and ± 7 days) to test whether relaxed alignment criteria could yield multivariate sampling states suitable for data-driven analysis. These windows were selected to reflect operationally realistic alignment tolerances within regulatory sampling schedules. No temporal averaging or resampling was applied beyond these matching windows, in order to preserve the integrity of the original monitoring records and avoid artificially inflating coherence.

Temporal coherence analysis was conducted independently of predictive modelling in order to diagnose structural sampling alignment prior to any reconstruction attempts. This separation ensures that observed limitations in multivariate completeness reflect inherent monitoring structure rather than modelling artefacts.

2.7. Multivariate reconstruction framework

Random Forest regression was implemented as a diagnostic stress test of multivariate coherence within the monitoring archive, rather than as a predictive optimisation exercise. The objective was not to maximise predictive accuracy but to determine whether the monitoring archive contains sufficient temporal coherence and multivariate completeness to support transferable sediment reconstruction across space. Random Forests were selected due to their ability to capture nonlinear relationships and interaction effects without parametric assumptions, while remaining robust to multicollinearity and irregular predictor distribution (Breiman, 2001; James et al., 2017).

RF is widely used in environmental monitoring contexts where predictor distributions are uneven and relationships are non-linear (Belgiu and Drăguţ, 2016). A simple multiple linear regression benchmark was also evaluated during preliminary testing to confirm that observed performance limitations were not model-class specific; Random Forest was retained as the primary framework due to its greater stability under grouped cross-validation. Results from the linear benchmark are not presented in detail, as performance patterns were consistent with those observed under the Random Forest framework.

Model configuration was intentionally conservative to reduce overfitting under sparse multivariate conditions. Hyperparameters were selected following preliminary grid-search exploration to balance model flexibility and generalisation stability (n_estimators = 500; min_samples_leaf = 2; max_depth unrestricted). Increasing tree depth beyond this configuration did not materially improve cross-validated performance but increased variance between folds and was therefore not adopted.

All models were trained using a fixed random seed to ensure reproducibility of validation results.

2.7.1. Predictor and response selection

Sediment indicators included suspended solids (105 °C), turbidity, and in situ turbidity. Candidate predictor variables were restricted to frequently monitored oxygen-related and physical parameters, including biochemical oxygen demand (BOD, 5-day ATU), chemical oxygen demand (COD), and water temperature.

Restriction to routinely monitored predictors ensures that reconstruction feasibility is evaluated under realistic regulatory sampling conditions rather than idealised variable inclusion.

Stations were considered eligible for reconstruction analysis if ≥60 sediment observations were recorded during the study period, ensuring a minimum level of temporal representation suitable for diagnostic modelling.

BOD and temperature were identified as the only routinely monitored variables that co-occurred with sediment observations within all tested temporal tolerance windows. COD, salinity and nutrient variables were excluded because multivariate completeness screening (Section 2.7.3) confirmed that no station–time event contained simultaneous observations of BOD, COD, and temperature within any matching tolerance. The restriction to empirically co-occurring predictors ensures that reconstruction feasibility is evaluated under realistic regulatory sampling conditions rather than idealised variable sets.

Beyond empirical co-occurrence, the selection of BOD and water temperature as candidate predictors is strongly supported by estuarine process theory. Water temperature directly governs fluid viscosity and localized hydrodynamic stratification, which dictate particulate settling velocities and flocculation dynamics in macrotidal estuaries (Geyer and MacCready, 2014; Winterwerp, 2002). Concurrently, BOD serves as a critical conceptual proxy for organic-matter loading; in historically industrialised channels, sediment resuspension events remobilise organic-rich benthic substrates, directly accelerating localized microbial respiration and biochemical oxygen demand within the water column (Moriarty et al., 2021; Spieckermann et al., 2022). Thus, tracking these variables concurrently evaluates the broader coupling between physical sediment dynamics and biogeochemical oxygen consumption under realistic regulatory sampling conditions rather than idealised, structurally unavailable variable sets.

2.7.2. Temporal alignment for modelling

Sediment observations were temporally paired with candidate predictors using symmetric tolerance windows (±0, ±3, ±7 days). For each sediment sampling event, the nearest routine observation within the specified window was selected.

If no routine measurement fell within the tolerance window, the sediment event was excluded from pairing.

2.7.3. Multivariate completeness assessment

Paired events were reshaped into a wide-format data matrix in which each row represented a station–time event and columns represented predictor variables and the sediment response variable.

Multivariate completeness was explicitly evaluated by quantifying the co-occurrence of candidate predictors within each paired event. Across all evaluated temporal windows, no station–time event contained simultaneous observations of BOD, COD, and temperature within the same matching tolerance. Consequently, reconstruction modelling was restricted to empirically co-occurring predictors (BOD and temperature), reflecting structural sampling constraints rather than theoretical exclusion of variables.

Rows containing missing values for the selected predictor set were excluded from modelling.

2.7.4. Regional pooling

Initial single-station modelling was constrained by sparse multivariate completeness. A pooled regional dataset was constructed across eligible stations while retaining station identifiers.

Pooling increases statistical power but does not assume spatial homogeneity; this assumption is explicitly tested through grouped cross-validation.

2.7.5. Model training and validation

Model performance was assessed using:

- 70/30 random train-test split
- Five-fold cross-validation
- Grouped cross-validation (GroupKFold) by station

The grouped cross-validation approach ensures that training and testing sets are spatially independent, thereby evaluating the capacity of reconstructed relationships to generalise across monitoring locations. The difference between performance under random splitting and grouped cross-validation is interpreted as a Transferability Gap (ΔR^2_transfer), representing the loss of predictive capacity when spatial dependence is removed. This gap provides a quantitative indicator of structural generalisability within the monitoring network.

Model performance metrics included coefficient of determination (R^2) and root mean square error (RMSE). Feature importance was extracted from the trained RF model to assess relative predictor contributions.

No feature scaling was applied, as tree-based ensemble algorithms are invariant to monotonic transformations of predictor variables.

Together, this hierarchical validation structure allows reconstruction performance to be evaluated under progressively stricter generalisation constraints, distinguishing within-station fit from cross-station transferability.

The overall analytical workflow, from raw regulatory archive extraction through to hierarchical validation of reconstruction performance, is summarised in Fig. 2. This framework distinguishes structural data diagnostics (classification, eligibility filtering, temporal matching, and completeness assessment) from predictive modelling and validation, emphasising the progressive narrowing of the dataset prior to reconstruction analysis.

2.8. Computational environment and reproducibility

All data processing, analysis, and visualization were conducted using Python (v3.x) within the Google Colab cloud computing environment. Standard open-source libraries were used for data handling, statistical analysis, and machine learning, consistent with reproducible environmental data analysis practices (James et al., 2017; Raschka and Mirjalili, 2019).

All scripts operate exclusively on openly available data from the Environment Agency archive, ensuring that the analysis can be reproduced or extended without access restrictions (Environment Agency, 2025).

3. Results

3.1. Overview of data availability and analytical subset

A total of 83,383 observations across 97 monitoring stations were initially extracted from the Environment Agency Water Quality Archive for the period January 2015 to January 2025. These stations represent the full regulatory monitoring footprint within the River Mersey estuarine and adjacent coastal area.

Following restriction to numeric water-column measurements with valid continuous results, and exclusion of non-water sample types (e.g., sediment cores, biota, effluents), the analytical dataset comprised 44,516 observations across 42 of the 97 stations and 187 distinct determinants (Table 1). This filtered subset represents the operational water-column monitoring architecture relevant to routine estuarine assessment and forms the basis for all subsequent frequency classification, process-based grouping, temporal coherence analysis, and reconstruction diagnostics.

To ensure structural interpretability of sediment monitoring within the network, progressive filtering was applied. Of the 42 water-column stations, 28 recorded at least one sediment- or turbidity-related measurement during the study period. Application of a ≥ 60 sediment-observation threshold, corresponding approximately to six

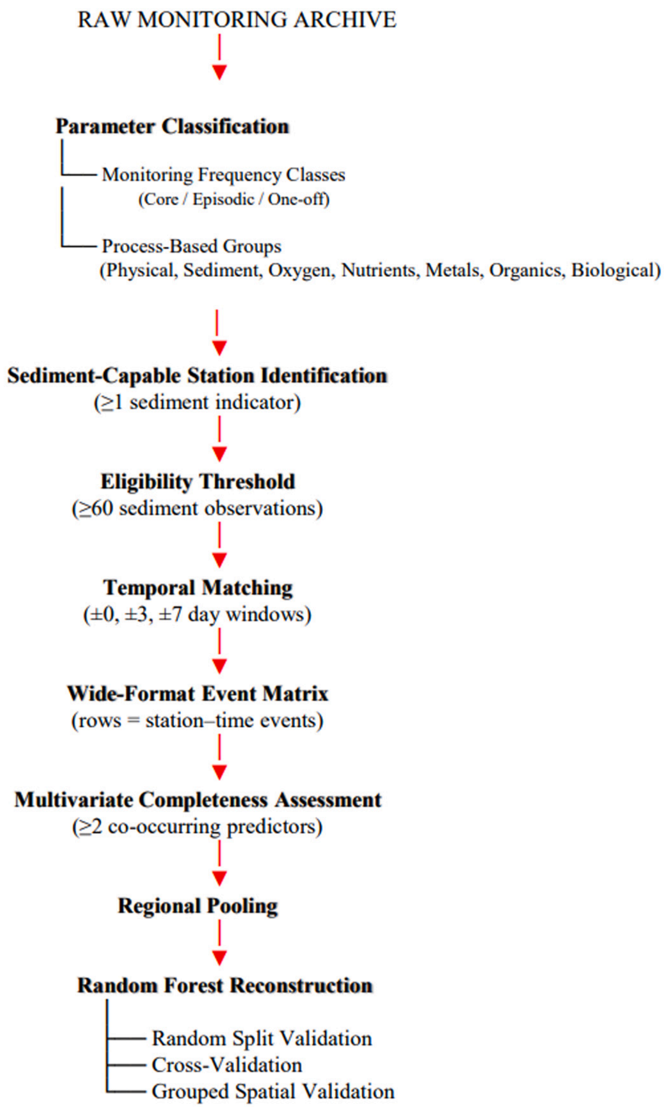

Fig. 2. Analytical workflow for multivariate sediment reconstruction from regulatory monitoring data. The diagram summarises the sequential data processing stages from archive extraction and dual parameter classification through eligibility filtering, temporal alignment, multivariate completeness assessment, and regional pooling, culminating in Random Forest modelling evaluated under hierarchical validation schemes. The workflow represents methodological design rather than quantitative outcomes.

Table 1
Summary of analytical dataset (water-column subset, 2015–2025).

Metric	Value
Total observations	44,516
Monitoring stations	42
Determinants	187
Core determinants (≥1000 obs)	13
Episodic determinants (100–1000 obs)	48
One-off determinants (<100 obs)	126
Stations with ≥1 sediment obs	28
Stations with ≥60 sediment obs	13
Final modelling rows	528
Stations used in RF modelling	6

observations per year over the study period, reduced the sediment-capable subset to 13 stations. This filtering step was implemented to retain stations with sufficient temporal continuity and co-sampling potential for subsequent multivariate coherence and reconstruction analyses.

Subsequent temporal alignment and multivariate completeness filtering further reduced the modelling-eligible subset to 6 stations, yielding 528 paired station–time events for reconstruction analysis. This sequential contraction from 42 stations to 6 modelling-eligible stations illustrates the structural narrowing imposed by sediment sparsity and multivariate misalignment within the regulatory archive.

This staged reduction reflects structural characteristics of the monitoring archive rather than data-processing exclusion and provides a transparent account of how sediment reconstruction feasibility narrows under realistic regulatory sampling conditions.

The total number of samples collected at each station over the ten-year period is summarised in Supplementary Material S2.

3.2. Frequency structure of water-column monitoring

Frequency classification of the 187 determinants revealed a highly skewed monitoring structure. Only 13 determinants (6.9%) met the ≥1000 observation threshold and were classified as core variables. A further 48 determinants (25.7%) were classified as episodic (100–1000 observations), while 126 determinants (67.4%) were recorded fewer than 100 times over the ten-year period and were therefore classified as one-off.

Despite the diversity of measured parameters, sustained regulatory surveillance is concentrated within a small core subset. Most determinants are measured intermittently or within targeted campaigns rather than under continuous long-term observation. As a result, the probability of constructing temporally coherent multivariate states declines rapidly outside the core monitoring subset.

3.3. Process-based composition of frequently monitored parameters

Analysis of the core and episodic determinants (≥100 observations) revealed a pronounced process-level bias in routine monitoring. Frequently monitored variables were dominated by physical and hydromorphological descriptors, including temperature, salinity, water depth, sampling depth, and tidal timing parameters. Oxygen-related metrics, including dissolved oxygen concentrations and saturation percentages, constituted the second most consistently monitored group.

When aggregated across the full water-column dataset ($n = 44{,}516$), physical and hydromorphological variables accounted for 23,138 observations (52.0%). Nutrients and productivity indicators contributed 7771 observations (17.5%), followed by oxygen and organic pollution metrics (4646; 10.4%), metals and inorganic contaminants (4223; 9.5%), and organic micropollutants (2884; 6.5%). In contrast, sediment and turbidity indicators represented only 1854 observations (4.2%), despite the estuary's strong tidal forcing and sediment transport dynamics.

Although sediment indicators were spatially distributed, 28 of 42 stations recorded at least one sediment-related measurement, sampling intensity per station remained substantially lower than that of hydrographic and oxygen-related variables. Sediment monitoring was therefore structurally present but operationally sparse.

Parameters associated with nutrients, metals, organic micropollutants, and sediment chemistry were predominantly classified as episodic or one-off, reflecting targeted investigative or compliance-driven campaigns rather than continuous surveillance. This imbalance indicates that parameters most directly associated with contaminant transport and sediment-bound pollutant dynamics are not embedded within sustained multivariate surveillance regimes.

Table 2 summarises the most frequently monitored water-column determinants, including sampling intensity, spatial coverage, and process grouping, while the complete classification is provided in Supplementary Material S3.

Table 2
Core and episodic water-column determinants (≥100 observations), in the River Mersey estuary (2015–2025), showing sampling frequency, spatial coverage, temporal variability, and process grouping.

Parameter	Unit	Total samples	Stations reporting	Process group
Temperature of Water	°C	2113	40	Physical & Hydromorphological
Oxygen, Dissolved (% Saturation)	%	2100	38	Oxygen & Organic Pollution
Oxygen, Dissolved as O_2	mg L^{-1}	2075	37	Oxygen & Organic Pollution
Salinity (In Situ)	‰	1819	30	Physical & Hydromorphological
Sample Depth below surface	m	1688	29	Physical & Hydromorphological
Water Depth	m	1578	28	Physical & Hydromorphological
Time of high tide	HH.MM	1538	24	Physical & Hydromorphological
Time relative to previous high water	HH.MM	1537	23	Physical & Hydromorphological
Turbidity (In Situ)	FTU	1323	23	Sediment & Turbidity
Volume of Sample Filtered	mL	1036	22	Physical & Hydromorphological
Chlorophyll (Acetone Extract)	μg L^{-1}	1008	22	Nutrients & Productivity
pH	–	826	32	Physical & Hydromorphological
Nitrite (Filtered as N)	mg L^{-1}	762	11	Nutrients & Productivity
Nitrate (Filtered as N)	mg L^{-1}	762	11	Nutrients & Productivity
Orthophosphate (Filtered as P)	mg L^{-1}	762	11	Nutrients & Productivity
Ammoniacal Nitrogen (Filtered as N)	mg L^{-1}	762	11	Nutrients & Productivity

3.4. *Distribution of monitoring effort over time*

Temporal aggregation by process group (Fig. 3) demonstrates sustained dominance of physical and hydromorphological measurements throughout the 2015–2025 period. Sediment and turbidity indicators remained consistently under-represented across all years. No sustained increase in sediment monitoring intensity was observed during high-disturbance periods, including the COVID-19 disruption interval.

Importantly, the lower total number of sediment observations does not primarily reflect absence of spatial representation. Rather, sediment sampling frequency per station was low relative to routine hydrographic monitoring. Sediment capture occurred at a majority of water-column stations at least once, but rarely at sustained or high-frequency intervals (Fig. 4). This pattern indicates that sediment monitoring is structurally distributed but temporally intermittent.

Red markers indicate stations where sediment or turbidity parameters were recorded at least once during the study period; grey markers indicate stations without sediment monitoring. Although sediment-related observations were numerically under-represented in the archive, spatial coverage was not limited to a small number of locations. Sediment measurements occurred at the majority of water-column stations at least once. The lower overall sediment sample count therefore reflects low sampling frequency per station, rather than absence of spatial representation, with sediment observations typically collected intermittently rather than under sustained routine monitoring.

3.5. *Feasibility of machine-learning gap-filling for sediment indicators*

The structural under-representation and temporal fragmentation of sediment indicators motivated a diagnostic assessment of reconstructability using routinely monitored predictors. Of the 42 water-column stations, 28 recorded at least one sediment observation and 13 recorded ≥60 sediment measurements. After temporal matching and multivariate completeness filtering, 528 paired station–time events across 6 stations were available for modelling.

Random Forest regression trained under random train–test splitting yielded moderate predictive performance (Test $R^2 = 0.588$; RMSE = 8.29). However, when evaluated using grouped cross-validation by station, a stricter test of spatial generalisation, mean R^2 declined to 0.258, indicating reduced transferability across monitoring locations. Grouped validation scores exhibited substantial inter-fold variability, reflecting heterogeneous station-level relationships and reinforcing the limited spatial consistency of sediment–hydrographic coupling across the network.

Model performance varied substantially by station (Table 3). In-sample R^2 exceeded 0.8 at several locations, suggesting that locally consistent relationships exist between sediment indicators and routine hydrographic variables. However, these relationships were not stable across stations. This contrast between within-station interpolation and cross-station extrapolation demonstrates that sediment–hydrographic coupling is predominantly site-specific rather than governed by transferable system-wide structure.

Feature importance analysis indicated that temperature accounted for approximately 64% of explained variance, with BOD contributing 36%, suggesting that model skill primarily reflects indirect hydrographic coupling rather than robust sediment-process representation. Impurity-based feature importance scores were strongly dominated by temperature (≈64%), with BOD contributing the remaining ≈36%. Because impurity-based importance measures are sensitive to predictor variance and correlation structure, these values should be interpreted as relative dominance within the fitted ensemble rather than as causal attribution or true variance partitioning. The observed importance pattern nevertheless indicates that predictive skill arises primarily from indirect hydrographic covariation rather than explicit sediment-process encoding.

These results indicate that while local interpolation is feasible under favourable station-specific conditions, routine hydrographic monitoring does not encode sufficient multivariate structure to support transferable sediment reconstruction across the estuarine network. Within the evaluated modelling framework, structural sparsity and asynchronous sampling appear to constrain generalisable reconstruction more strongly than model flexibility.

Although only a Random Forest ensemble was implemented in this study, the pronounced decline in performance under grouped cross-validation reflects spatial non-stationarity and incomplete multivariate co-sampling rather than limitations inherent to a specific algorithmic architecture. Tree-based ensembles are generally robust to nonlinear structure and multicollinearity; therefore, the observed transferability gap is more parsimoniously attributed to structural data fragmentation than to model selection. While alternative regression frameworks could be explored in future work, increases in algorithmic complexity alone are unlikely to compensate for absent temporal alignment between sediment observations and routinely monitored predictors.

3.6. *Quantification of temporal separation and co-sampling coherence between monitoring programmes*

To investigate whether the loss of cross-station predictive performance observed in Section 3.5 arises from temporal sampling structure, a quantitative assessment of inter-programme temporal separation was conducted. Eighteen stations met the sediment-capable threshold and were included in the temporal separation analysis. For each sediment observation, the minimum temporal distance to the nearest sampling event from other process groups was calculated (Table 4).

Although median separations were zero days across all comparison groups, this statistical outcome does not indicate systematic or routine co-sampling across the network. Instead, it reflects a heavily skewed,

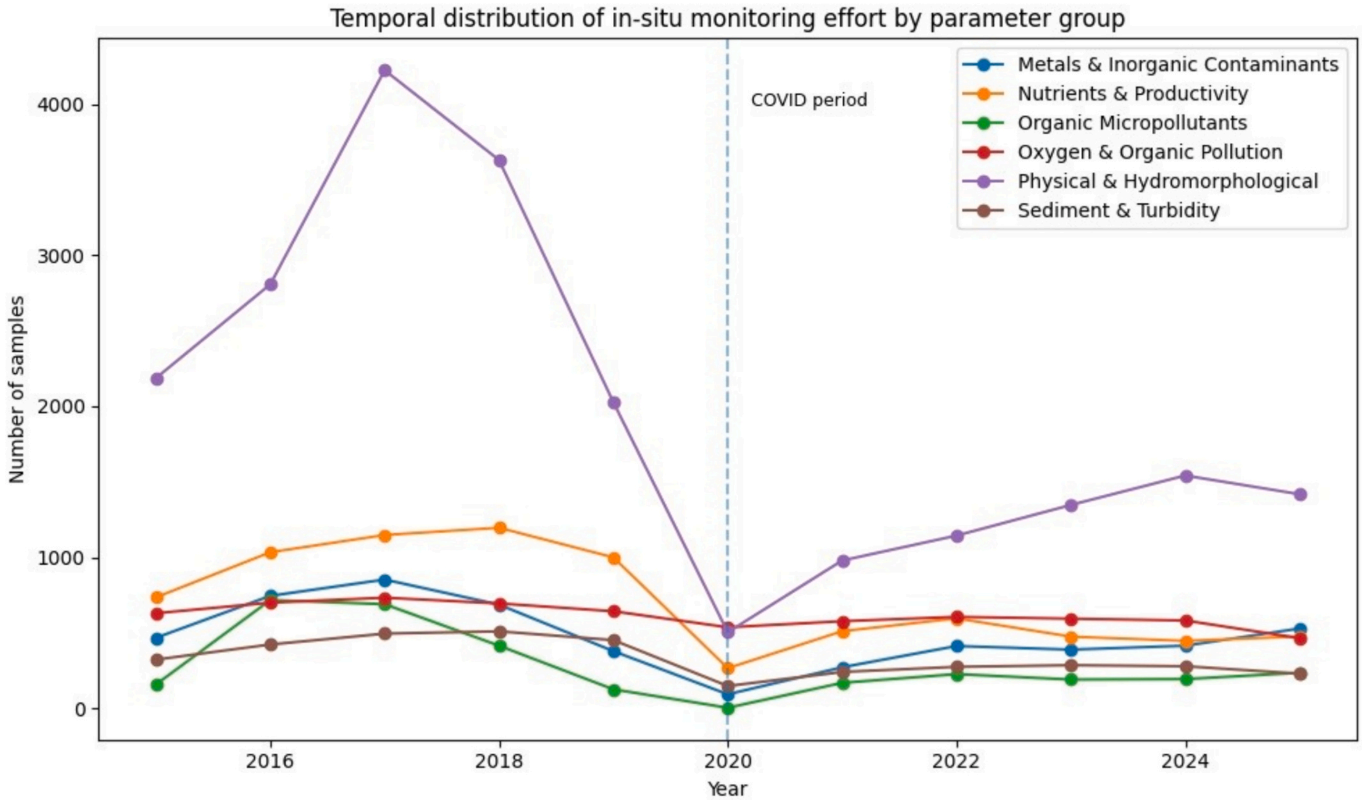

Fig. 3. Temporal evolution of in-situ monitoring effort by parameter group (2015–2025), showing persistent dominance of physical and hydromorphological measurements and sustained under-representation of sediment and turbidity indicators. The dashed line marks the COVID-19 period.

Fig. 4. Spatial distribution of sediment-capable and non-sediment monitoring stations within the River Mersey estuary (2015–2025).

bimodal temporal structure in which infrequent, episodic co-sampling campaigns produce zero-separation instances that are embedded within extended periods of asynchronous measurement.

The broader distributions reveal substantial temporal fragmentation. Mean separations ranged from approximately 46 days for metals and inorganic contaminants to nearly 380 days for nutrients and productivity indicators. Upper-tail separations were particularly pronounced, with 90th-percentile values exceeding 1800 days for both nutrient and physical variables and maximum separations extending over multiple years. Oxygen-related parameters exhibited comparatively smaller mean separations (≈118 days), yet even within this group, 10% of sediment observations were separated from the nearest oxygen

Table 3
Station-level in-sample random forest performance for sediment reconstruction.

Station ID	samples	R^2 (in-sample)
NW-88001072	23	0.958
NW-88002799	82	0.823
NW-88002787	96	0.820
NW-88002783	91	0.817
NW-88002778	95	0.725
NW-88002796	84	0.618

measurement by more than 230 days.

These results indicate that while coordinated sampling events do occur, sediment monitoring is not systematically embedded within routine hydrographic or biogeochemical programmes. The observed temporal fragmentation indicates that sediment observations were frequently collected under environmental conditions substantially different from those represented by associated hydrographic or contaminant measurements.

While temporal separation statistics describe timing differences between monitoring programmes, they do not directly quantify monitoring coherence. To evaluate the degree of temporal alignment between sediment observations and other routinely monitored variables, the Co-sampling Coherence Ratio (CCR) defined in Section 2.6 was applied.

Pairwise CCR analysis revealed substantial variation between process groups. Network-level CCR values were highest for Oxygen & Organic Pollution variables (CCR = 0.93), followed by Physical & Hydromorphological variables (CCR = 0.79), Nutrients & Productivity (CCR = 0.75), and Metals & Inorganic Contaminants (CCR = 0.67). In contrast, Organic Micropollutants exhibited markedly lower coherence (CCR = 0.27), indicating limited synchronisation with sediment observations despite occasional coordinated sampling events.

However, pairwise coherence alone does not guarantee the existence of complete multivariate observation states suitable for integrated analysis. Consequently, a stricter coherence metric (CCR_{strict}) was calculated requiring simultaneous availability of all variables included within the final multivariate reconstruction framework. Whereas CCR measures monitoring alignment between sediment observations and individual process groups, CCR_{strict} measures analytical readiness for multivariate reconstruction.

Across the 18 sediment-capable stations, mean CCR_{strict} was 0.28 and the median value was only 0.07, indicating that complete multivariate records were uncommon throughout the monitoring archive. Furthermore, 10 of the 18 sediment-capable stations exhibited CCR_{strict} values of zero, demonstrating a complete absence of fully aligned multivariate observations despite the presence of sediment monitoring. In contrast, a small subset of stations exhibited high CCR_{strict} values (>0.7), indicating that coordinated multivariate sampling had effectively occurred at these locations. These stations corresponded closely to the subset retained for machine-learning reconstruction (See Table 5.).

The contrast between CCR and CCR_{strict} is important. While pairwise co-sampling occurred intermittently across several process groups, analytically usable multivariate records remained concentrated within a limited subset of stations. These findings indicate that analytical usability is constrained not only by observation frequency, but also by the degree of multivariate temporal coordination embedded within the monitoring design. Detailed station-level coherence statistics, including process-specific CCR values and sampling requirements, are provided in Supplementary Material S4.

The combination of zero medians and long upper tails reveals a bimodal monitoring structure: episodic co-sampling events embedded within extended periods of temporal decoupling. When considered alongside the reduced cross-station predictive performance of the Random Forest model (mean grouped $R^2 \approx 0.26$), this temporal fragmentation provides a structural explanation for the limited transferability of sediment reconstruction across the monitoring network.

The observed ΔR^2_transfer of approximately 0.33 (0.588–0.258) indicates that more than half of the apparent predictive skill under random splitting is attributable to within-station structure rather than transferable system-wide relationships. This magnitude of performance collapse suggests that sediment reconstruction becomes unstable when multivariate coherence is spatially fragmented across monitoring locations.

The temporal diagnostics presented here therefore indicate that reconstructability is intrinsically linked to co-sampling architecture. Restricting analysis to strictly co-occurring observations would likely increase multivariate completeness while reducing sample size, highlighting the trade-off between data volume and structural coherence in regulatory archives. The observed fragmentation pattern indicates that transferable sediment reconstruction is constrained by monitoring architecture itself. Enhanced same-day or near-synchronous co-sampling of sediment and contaminant indicators would be expected to improve cross-station stability, reinforcing the conclusion that monitoring configuration rather than modelling flexibility governs pollution-detection capacity within the regulatory archive.

The observed temporal decoupling between sediment indicators and contaminant measurements has direct implications for pollution surveillance. In estuaries with contaminated legacy sediments, resuspension events can release metals and organic pollutants into the water column over short time scales. When sediment monitoring is

Table 5
Summary of pairwise monitoring coherence (CCR) and strict multivariate coherence (CCR_{strict}) across sediment-capable stations.

Metric	Value	Process Group	Network CCR
Sediment-capable stations	18	Oxygen & Organic Pollution	0.93
Mean CCR_{strict}	0.28	Physical & Hydromorphological	0.79
Median CCR_{strict}	0.07	Nutrients & Productivity	0.75
Stations with CCR_{strict} = 0	10	Metals & Inorganic Contaminants	0.67
Stations with CCR_{strict} > 0.7	6	Organic Micropollutants	0.27

Note: High CCR_{strict} values (>0.7) are reported descriptively to identify stations with near-complete multivariate alignment and do not represent a formal threshold.

Table 4
Descriptive statistics of the minimum temporal separation (in days) between each sediment observation and the nearest sampling event from other routinely monitored process groups, calculated across all sediment-capable stations.

Comparison group	N (sediment obs)	Mean (days)	25th pct	Median	75th pct	90th pct	Max (days)
Metals & Inorganic Contaminants	1053	46.35	0.0	0.0	0.04	68.56	1555.12
Nutrients & Productivity	1606	379.79	0.0	0.0	29.95	1803.64	3490.09
Organic Micropollutants	523	113.10	0.0	0.0	26.99	141.81	2332.06
Oxygen & Organic Pollution	1308	117.67	0.0	0.0	0.002	236.87	2244.97
Physical & Hydromorphological	1772	370.68	0.0	0.0	0.0	1782.91	3333.03

Note: Zero-day separations reflect occasional same-day or near-simultaneous sampling during coordinated monitoring campaigns. Large mean values and long upper tails indicate prolonged periods of temporal decoupling between sediment observations and other monitoring programmes.

asynchronous with chemical measurements, the risk of failing to resolve these episodic mobilisation events may increase within regulatory archives, creating conditions under which sediment-associated contaminant episodes may remain incompletely characterised within routine regulatory archives.

4. Discussion

4.1. Structural configuration of sediment monitoring within the regulatory network

The monitoring archive exhibited a strongly uneven sampling structure, with routine effort concentrated within a relatively small subset of physical and oxygen-related variables, while sediment indicators remained comparatively sparse. Although sediment observations were geographically distributed across much of the estuary, they were typically collected intermittently rather than through sustained monitoring programmes. Such configurations are characteristic of compliance-oriented monitoring networks, where sediment variables are often embedded within episodic investigations rather than coordinated process-based surveillance (Defeo and Elliott, 2021).

Temporal separation statistics provide quantitative confirmation of this fragmentation. Although median separations between sediment and other process groups were frequently zero days, indicating that coordinated campaigns occasionally occurred, the broader distributions revealed prolonged asynchronous sampling. Mean separations ranged from 46 days (metals) to nearly 380 days (nutrients and physical variables), with upper-percentile separations extending over multiple years. Even oxygen-related variables exhibited mean separations exceeding 100 days.

Importantly, zero-day median separations do not indicate routine co-sampling across the network. Rather, they reflect episodic coordinated campaigns embedded within substantially longer periods of asynchronous measurement. This bimodal structure is environmentally significant because sediment-associated contaminants may shift rapidly between particulate and dissolved phases during tidal resuspension and estuarine mixing processes operating over tidal to sub-daily timescales (Geyer et al., 2001; Robins et al., 2016; Turner and Millward, 2002). Consequently, contaminant measurements collected days or weeks apart from sediment observations may not represent the same transient mobilisation state. This is particularly relevant in historically industrialised estuaries such as the Mersey, where legacy contaminants stored within fine sediments may be episodically remobilised during energetic hydrodynamic conditions (Eggleton and Thomas, 2004; Hawkins et al., 2020; Langston et al., 2006).

Turbidity alone therefore provides only a partial indication of sediment-mediated contaminant mobilisation unless sediment composition and particulate characteristics are also considered. Sediment texture, organic matter content, and particle-size distribution may strongly influence adsorption-desorption behaviour and transport dynamics within estuarine systems (Eggleton and Thomas, 2004; Turner and Millward, 2002). However, these descriptors were not systematically integrated within the analysed archive, limiting interpretability even where temporal alignment existed. These processes may also influence contaminant bioavailability and ecological exposure pathways for aquatic organisms, further emphasising the importance of coordinated sediment and contaminant monitoring within ecosystem-based management frameworks.

In addition to temporal fragmentation, sediment-associated contaminant transport is influenced by strong spatial and physicochemical gradients within estuarine environments. Many contaminants behave non-conservatively within mixing zones because of salinity-dependent partitioning, particle interactions, and early diagenetic processes (Turner and Millward, 2002). In highly turbid estuaries, sediment transport pathways may also shift spatially through migration of estuarine turbidity maximum (ETM) zones, producing transient hotspots that fixed monitoring stations may not consistently resolve (Geyer et al., 2001).

These findings suggest that limitations are linked more strongly to monitoring configuration and temporal coherence than to overall data volume alone. Although pairwise CCR values indicated moderate-to-high coherence for several monitoring domains, strict multivariate coherence (CCR_{strict}) remained low across most sediment-capable stations. Apparent co-sampling therefore did not necessarily translate into complete analytical states suitable for integrated sediment–contaminant assessment. Station-level diagnostics (Supplementary Material S4) further demonstrated substantial heterogeneity in multivariate alignment across the monitoring network.

Although the River Mersey represents a historically industrialised macrotidal estuary with strong tidal resuspension dynamics, the monitoring limitations identified here are unlikely to be unique to this system. Many long-established regulatory monitoring programmes were originally designed around compliance-focused chemical surveillance rather than coordinated process-based sediment monitoring (Elliott and Whitfield, 2011). Consequently, asynchronous sampling between sediment, hydrographic, and contaminant variables may also occur in other estuarine and coastal monitoring archives. Nevertheless, the magnitude of temporal fragmentation, CCR- and CCR_{strict}-based diagnostics, and reconstruction limitations observed here are likely to remain system-dependent and influenced by local hydrodynamic conditions, estuarine morphology, contaminant history, and monitoring objectives. Cross-system evaluation across contrasting estuarine environments is therefore required before the proposed coherence thresholds can be considered broadly transferable.

4.2. Diagnostic evaluation of reconstructability

Random Forest regression was applied as a diagnostic tool to evaluate whether routinely monitored hydrographic variables contain sufficient information to reconstruct sediment indicators. Ensemble tree-based methods are well established for modelling nonlinear environmental relationships without restrictive parametric assumptions (Belgiu and Drăguţ, 2016; Tyralis et al., 2019).

Predictive performance was moderate under random train–test splitting ($R^2 = 0.588$) but declined substantially under grouped cross-validation ($R^2 = 0.258$), indicating limited transferability across stations. Because random splitting evaluates interpolation within stations whereas grouped cross-validation assesses transferability across spatially independent locations (Roberts et al., 2017), the observed performance gap ($\Delta R2_{transfer} \approx 0.33$) suggests that sediment–hydrographic relationships were largely site-specific rather than network-consistent.

Temporal misalignment appears to represent the dominant limitation affecting reconstruction. However, spatial heterogeneity across estuarine zones also contributed to the observed transferability gap. The modelling stations spanned contrasting hydrodynamic environments, ranging from inner-estuary regions influenced by tidal forcing and freshwater discharge to outer-estuary zones characterised by differing salinity and wave-energy regimes (Geyer and MacCready, 2014). Consequently, sediment-hydrographic relationships likely reflected local environmental conditions rather than a single transferable system-wide relationship.

This interpretation is consistent with the observed coherence structure, where differences in multivariate alignment (CCR_{strict}) result in locally stable relationships that do not generalise across stations. Coherence diagnostics further reveal a critical distinction between apparent and functional monitoring alignment. While pairwise CCR indicates that sediment observations often coincide with at least one process variable, strict multivariate coherence (CCR_{strict}) demonstrates that complete observation states remain limited to a small subset of stations.

Feature-importance analysis indicated dominance of temperature, with secondary contribution from BOD. Although Random Forest importance metrics may be influenced by predictor covariance (Strobl et al., 2007), this pattern suggests that predictive performance primarily reflects indirect hydrographic relationships rather than explicit representation of sediment mobilisation processes. Temperature and BOD were retained because they represented the most consistently co-occurring hydrographic variables across sediment-capable stations, whereas several chemically relevant variables exhibited insufficient temporal overlap for stable multivariate reconstruction.

These findings indicate that routine hydrographic monitoring does not encode sediment variability in a transferable, system-level manner across the estuarine network. The primary limitation therefore appears to arise from monitoring design, temporal coherence, and environmental heterogeneity rather than modelling complexity alone.

Although only a Random Forest ensemble was implemented during this diagnostic stage, the marked decline in grouped-validation performance likely reflects environmental heterogeneity and incomplete multivariate co-sampling rather than limitations specific to a single modelling algorithm. Alternative machine-learning approaches, including Gradient Boosting Regressors or K-Nearest Neighbours, may alter absolute performance metrics to some extent, but are unlikely to compensate fully for structural sparsity and asynchronous sampling embedded within the monitoring design.

These findings also have practical implications for future monitoring-network design. Rather than uniformly increasing sampling volume, greater analytical benefit may arise from improving temporal coordination between sediment, hydrographic, and contaminant observations. Specific monitoring recommendations are discussed further in Section 4.3.

4.3. Monitoring design implications and the role of satellite observations

One potential mechanism for improving temporal responsiveness within fragmented monitoring systems involves integration with higher-frequency observational platforms.

Satellite remote sensing represents a fundamentally different observational configuration. Multispectral platforms such as European Space Agency Sentinel-2 provide spatially continuous surface reflectance measurements at regular revisit intervals, enabling repeated system-wide observation of turbidity-related optical properties (Dogliotti et al., 2015; Salama et al., 2022). In the lower River Mersey, recent work has demonstrated that when satellite imagery is temporally aligned with in-situ sampling, ensemble machine-learning models can achieve high predictive performance ($R^2 = 0.84$) for turbidity estimation (Nangir et al., 2025).

However, the applicability of satellite remote sensing within sediment-mediated pollution surveillance remains restricted to optically active surface properties, including turbidity, total suspended matter (TSM), and suspended particulate matter (SPM). Satellite observations cannot directly resolve dissolved metals, organic micropollutants, nutrient speciation, or sediment geochemical characteristics that govern contaminant fate and bioavailability within estuarine systems (Dogliotti et al., 2015; IOCCG, 2019; Salama et al., 2022). Consequently, remote sensing should be viewed not as a replacement for in-situ contaminant monitoring, but as a complementary observational layer capable of identifying elevated turbidity events and spatially dynamic sediment hotspots that may guide adaptive field sampling.

The relevance of satellite observations in the present context therefore lies primarily in their observational geometry rather than algorithmic superiority. Satellite systems provide spatially synchronous coverage at fixed revisit intervals, whereas the analysed in-situ network exhibited asynchronous and parameter-fragmented sampling. The contrast consequently reflects differences in monitoring architecture rather than modelling methodology alone.

The present study did not incorporate satellite-derived predictors within the Random Forest reconstruction framework and therefore does not provide a direct comparative evaluation. Nevertheless, the temporal-fragmentation patterns identified here suggest that observational systems capable of improving temporal consistency and spatial representativeness may help address weaknesses inherent in compliance-oriented monitoring networks.

The key implication therefore relates more strongly to monitoring architecture than to modelling sophistication, particularly the need for systematic co-sampling and multivariate alignment across key physical, chemical, and sediment domains.

From an operational perspective, improving sediment-mediated pollution surveillance may depend less on increasing total sampling volume and more on strengthening coordination between sediment, hydrographic, and contaminant observations. Based on the temporal-coherence diagnostics presented here, several practical redesign priorities emerge: (i) same-day co-sampling of sediment and contaminant variables at core stations; (ii) improved integration of turbidity and suspended-solids monitoring across routine programmes; (iii) adaptive event-triggered sampling during storms, dredging, or high-resuspension conditions; and (iv) prioritisation of ETM zones and recurrent sediment-retention hotspots for intensified monitoring.

Table 6 summarises a proposed tiered co-sampling framework derived from the diagnostic findings of this study.

This framework suggests that relatively modest improvements in temporal coordination may substantially improve sediment-associated pollution detectability without requiring complete redesign of existing regulatory infrastructure. Improving both temporal alignment and multivariate completeness increases the likelihood that sediment, contaminant, and hydrographic observations represent comparable environmental states suitable for integrated analysis.

4.4. Broader relevance

Although this analysis focuses on the River Mersey estuary, the structural characteristics identified are consistent with patterns documented across long-established regulatory monitoring networks. In many estuarine and coastal systems, monitoring programmes have historically prioritised compliance-driven indicators and baseline physical–chemical descriptors, while sediment-related variables are sampled less frequently or through targeted campaigns (Defeo and Elliott, 2021).

Such configurations largely reflect compliance-oriented regulatory mandates rather than integrated process-based system analysis. Comparable industrialised estuaries, including the Scheldt Estuary and Sydney Harbour, have similarly reported strong physicochemical variability, sediment-associated contamination pressures, and complex

Table 6
Proposed tiered co-sampling framework for sediment-mediated pollution surveillance in estuarine monitoring networks.

Tier	Monitoring objective	Sampling approach	Key variables	Suggested trigger/ frequency
Tier 1	Improve routine multivariate coherence	Same-day co-sampling at core stations	Turbidity, suspended solids, sediment indicators, dissolved oxygen, salinity, contaminants	Routine regulatory sampling
Tier 2	Resolve seasonal sediment dynamics	Coordinated seasonal or monthly surveys	Sediment composition, hydrographic variables, contaminant panels	Monthly or seasonal surveys
Tier 3	Capture episodic mobilisation events	Adaptive event-triggered sampling	Turbidity anomalies, dissolved and particulate contaminants	Storms, spring tides, dredging, high-flow events

sediment transport dynamics that challenge long-term environmental surveillance (Fettweis et al., 1998; Mayer-Pinto et al., 2015).

Although the absolute temporal separation values and reconstruction performance observed here are likely system-specific, the broader finding that multivariate temporal coherence constrains sediment-associated pollution detectability may be transferable to other compliance-driven monitoring networks.

For regulatory agencies operating under frameworks such as the Water Framework Directive or comparable environmental quality standards, this mismatch between observational design and analytical ambition may limit the capacity to evaluate sediment resilience, contaminant pathways, and coupled sediment-water interactions on a system scale (Borja et al., 2010). The diagnostic framework developed here, combining frequency distribution analysis, temporal separation metrics, pairwise coherence diagnostics (CCR), strict multivariate coherence diagnostics (CCR_{strict}), and hierarchical cross-validation, provides a structured approach for evaluating whether long-term monitoring archives are capable of resolving sediment–water interactions and transferable contaminant dynamics within complex estuarine systems.

As environmental management priorities increasingly shift toward ecosystem-based assessment, contaminant transport pathways, and coupled sediment–water interactions, monitoring architectures designed primarily for compliance may become structurally misaligned with process-based analytical requirements (Borja et al., 2010).

Cross-system comparative evaluation was outside the scope of the present study, but represents an important next step for assessing the transferability of CCR- and CCR_{strict}-based diagnostics and temporal coherence analysis across estuarine systems characterised by differing tidal regimes, sediment dynamics, and regulatory monitoring structures.

5. Conclusions

This study demonstrates that the capacity of long-term regulatory monitoring archives to support sediment-mediated contaminant assessment is strongly influenced by monitoring architecture rather than data volume alone. Although the River Mersey archive contained extensive water-column observations, sediment indicators remained temporally fragmented and inconsistently aligned with hydrographic and contaminant variables. While pairwise co-sampling coherence was often moderate to high, strict multivariate coherence remained low across most of the monitoring network, indicating that complete analytical states suitable for integrated assessment were uncommon.

Diagnostic machine-learning evaluation revealed a substantial transferability gap between within-station interpolation and cross-station generalisation, indicating that apparent predictive performance under random validation primarily reflected local structure rather than transferable system-wide sediment-process relationships. Reduced grouped-validation performance likely reflected the combined influence of asynchronous sampling, incomplete multivariate alignment, and spatial environmental heterogeneity across the estuarine system.

The findings suggest that improving sediment-mediated contaminant assessment in dynamic estuarine environments may depend less on increasing model complexity and more on strengthening temporal co-ordination between sediment, hydrographic, and contaminant monitoring programmes. Coordinated co-sampling, particularly during periods of elevated turbidity and sediment resuspension, is likely to provide greater analytical benefit than increasing sampling volume alone.

Although the present study did not directly quantify missed pollution events, it introduces a diagnostic framework based on temporal separation analysis, pairwise coherence (CCR), strict multivariate coherence (CCR_{strict}), and hierarchical validation. Together, these tools provide a practical basis for evaluating whether monitoring networks are structurally capable of resolving sediment-mediated contaminant dynamics and may offer a transferable framework for assessing other compliance-oriented estuarine monitoring systems.

CRediT authorship contribution statement

Deelaram Nangir: Writing – original draft, Visualization, Software, Methodology, Formal analysis, Data curation, Conceptualization. **Manolia Andredaki:** Writing – review & editing, Validation, Supervision, Conceptualization. **Iacopo Carnacina:** Writing – review & editing, Supervision, Methodology, Conceptualization.

Declaration of Generative AI and AI-assisted technologies in the writing process

The authors used an AI-based language tool to assist with grammar refinement and clarity improvement. All scientific content, analysis, and interpretations were developed and verified by the authors.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors thank the UK Environment Agency for providing open-access water quality monitoring data used in this study. The authors also acknowledge the use of open-source Python libraries and the Google Colab computational environment for data processing and analysis.

This research did not receive any specific grant from funding agencies in the public, commercial, or non-profit sectors.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.marpolbul.2026.120004.

Data availability

Data used in this study are publicly available from the UK Environment Agency Water Quality Archive (Water Quality Explorer).

References

Akhlaq, M., Ellahi, A., Niaz, R., Khan, M., Sammen, S.Sh., Scholz, M., 2024. Comparative Analysis of Machine Learning Algorithms for Water Quality Prediction. Tellus A: Dynamic Meteorology and Oceanography 76, 177–192. https://doi.org/10.16993/tellusa.4069.

Becker, A., Plater, A., Wolf, J., 2017. The Energy River: Realising Energy Potential from the River Mersey.

Behmel, S., Damour, M., Ludwig, R., Rodriguez, M.J., 2016. Water quality monitoring strategies — A review and future perspectives. Sci. Total Environ. 571, 1312–1329. https://doi.org/10.1016/j.scitotenv.2016.06.235.

Belgiu, M., Drăguţ, L., 2016. Random forest in remote sensing: A review of applications and future directions. ISPRS J. Photogramm. Remote Sens. 114, 24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011.

Bieroza, M.Z., Heathwaite, A.L., Bechmann, M., Kyllmar, K., Jordan, P., 2018. The concentration-discharge slope as a tool for water quality management. Sci. Total Environ. 630, 738–749. https://doi.org/10.1016/j.scitotenv.2018.02.256.

Borja, Á., Elliott, M., Carstensen, J., Heiskanen, A.-S., van de Bund, W., 2010. Marine management – Towards an integrated implementation of the European Marine Strategy Framework and the Water Framework Directives. Mar. Pollut. Bull. 60, 2175–2186. https://doi.org/10.1016/j.marpolbul.2010.09.026.

Breiman, L., 2001. Random Forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324.

Burton, 2003. The Mersey Basin: an historical assessment of water quality from an anecdotal perspective. Sci. Total Environ. 314–316, 53–66. https://doi.org/10.1016/S0048-9697(03)00094-9.

Chicco, D., Warrens, M.J., Jurman, G., 2021. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. PeerJ Comput. Sci. 7, e623. https://doi.org/10.7717/peerj-cs.623.

Chowdhury, M., Vilas, C., van Bergeijk, S., Navarro, G., Laiz, I., Caballero, I., 2023. Monitoring turbidity in a highly variable estuary using Sentinel 2-A/B for ecosystem management applications. Front. Mar. Sci. 10. https://doi.org/10.3389/fmars.2023.1186441.

Defeo, O., Elliott, M., 2021. The 'triple whammy' of coasts under threat – Why we should be worried! Mar. Pollut. Bull. 163, 111832. https://doi.org/10.1016/j.marpolbul.2020.111832.

Dogliotti, A.I., Ruddick, K.G., Nechad, B., Doxaran, D., Knaeps, E., 2015. A single algorithm to retrieve turbidity from remotely-sensed data in all coastal and estuarine waters. Remote Sens. Environ. 156, 157–168. https://doi.org/10.1016/j.rse.2014.09.020.

Eggleton, J., Thomas, K.V., 2004. A review of factors affecting the release and bioavailability of contaminants during sediment disturbance events. Environ. Int. 30, 973–980. https://doi.org/10.1016/j.envint.2004.03.001.

Elliott, M., Whitfield, A.K., 2011. Challenging paradigms in estuarine ecology and management. Estuar. Coast. Shelf Sci. 94, 306–314. https://doi.org/10.1016/j.ecss.2011.06.016.

Environment Agency, 2025. Water Quality Explorer [WWW Document]. Department for Food Environment and Rural Affairs, URL. https://environment.data.gov.uk/water-quality. accessed 1.9.26.

Fettweis, M., Sas, M., Monbaliu, J., 1998. Seasonal, Neap-spring and Tidal Variation of Cohesive Sediment Concentration in the Scheldt Estuary. Belgium. Estuar. Coast. Shelf Sci. 47, 21–36. https://doi.org/10.1006/ecss.1998.0338.

Geyer, W.R., MacCready, P., 2014. The Estuarine Circulation. Annu. Rev. Fluid Mech. 46, 175–197. https://doi.org/10.1146/annurev-fluid-010313-141302.

Geyer, W.R., Woodruff, J.D., Traykovski, P., 2001. Sediment Transport and Trapping in the Hudson River Estuary. Estuaries 24, 670. https://doi.org/10.2307/1352875.

Grandjean, T.J., Weenink, R., van der Wal, D., Addink, E.A., Hu, Z., Liu, S., Wang, Z.B., Yuan, L., Bouma, T.J., 2024. Critical turbidity thresholds for maintenance of estuarine tidal flats worldwide. Nat. Geosci. 17, 539–544. https://doi.org/10.1038/s41561-024-01431-3.

Greenwood, N., Devlin, M.J., Best, M., Fronkova, L., Graves, C.A., Milligan, A., Barry, J., van Leeuwen, S.M., 2019. Utilizing Eutrophication Assessment Directives From Transitional to Marine Systems in the Thames Estuary and Liverpool Bay. UK. Front. Mar. Sci. 6. https://doi.org/10.3389/fmars.2019.00116.

Hawkins, S.J., O'Shaughnessy, K.A., Adams, L.A., Langston, W.J., Bray, S., Allen, J.R., Wilkinson, S., Bohn, K., Mieszkowska, N., Firth, L.B., 2020. Recovery of an urbanised estuary: Clean-up, de-industrialisation and restoration of redundant dock-basins in the Mersey. Mar. Pollut. Bull. 156, 111150. https://doi.org/10.1016/j.marpolbul.2020.111150.

Hurley, R.R., Rothwell, J.J., Woodward, J.C., 2017. Metal contamination of bed sediments in the Irwell and Upper Mersey catchments, northwest England: exploring the legacy of industry and urban growth. J. Soils Sediments 17, 2648–2665. https://doi.org/10.1007/s11368-017-1668-6.

James, Gareth, Witten, Daniela, Hastie, Trevor, Tibshirani, Robert, 2017. An introduction to statistical learning : with applications in R. Springer, Springer Science+Business Media.

Karthick, M., Shanmugam, P., Saravana Kumar, G., 2024. Long-term water quality assessment in coastal and inland waters: An ensemble machine-learning approach using satellite data. Mar. Pollut. Bull. 209, 117036. https://doi.org/10.1016/j.marpolbul.2024.117036.

Kumar, L., Afzal, M.S., Ahmad, A., 2022. Prediction of water turbidity in a marine environment using machine learning: A case study of Hong Kong. Reg. Stud. Mar. Sci. 52, 102260. https://doi.org/10.1016/j.rsma.2022.102260.

Langston, W.J., Chesman, B.S., Burt, G.R., 2006. Site Characterisation of European Marine Sites The Mersey Estuary SPA.

Liu, Y., Song, Y., Liu, P., Lai, L., Zhang, Y., 2025. Concerns about the temporal matching windows in satellite-ground synchronization for lacustrine environment mapping. Water Res. 286, 124208. https://doi.org/10.1016/j.watres.2025.124208.

Lonsdale, J.-A., Leach, C., Parsons, D., Barkwith, A., Manson, S., Elliott, M., 2022. Managing estuaries under a changing climate: A case study of the Humber Estuary. UK. Environ. Sci. Policy 134, 75–84. https://doi.org/10.1016/j.envsci.2022.04.001.

Lunt, J., Smee, D.L., 2020. Turbidity alters estuarine biodiversity and species composition. ICES J. Mar. Sci. 77, 379–387. https://doi.org/10.1093/icesjms/fsz214.

Magrì, S., Ottaviani, E., Prampolini, E., Besio, G., Fabiano, B., Federici, B., 2023. Application of machine learning techniques to derive sea water turbidity from Sentinel-2 imagery. Remote Sens. Appl. 30, 100951. https://doi.org/10.1016/j.rsase.2023.100951.

Matos, T., Martins, M.S., Henriques, R., Goncalves, L.M., 2024. A review of methods and instruments to monitor turbidity and suspended sediment concentration. J. Water Process Eng. https://doi.org/10.1016/j.jwpe.2024.105624.

Mayer-Pinto, M., Johnston, E.L., Hutchings, P.A., Marzinelli, E.M., Ahyong, S.T., Birch, G., Booth, D.J., Creese, R.G., Doblin, M.A., Figueira, W., Gribben, P.E., Pritchard, T., Roughan, M., Steinberg, P.D., Hedge, L.H., 2015. Sydney Harbour: a review of anthropogenic impacts on the biodiversity and ecosystem function of one of the world's largest natural harbours. Mar. Freshw. Res. 66, 1088–1105. https://doi.org/10.1071/MF15157.

IOCCG, 2019. Uncertainties in ocean Colour remote sensing. In: Mélin, F. (Ed.), IOCCG Report Series, No. 18. International Ocean Colour Coordinating Group, Dartmouth, Canada. https://doi.org/10.25607/OBP696.

Moriarty, J.M., Friedrichs, M.A.M., Harris, C.K., 2021. Seabed Resuspension in the Chesapeake Bay: Implications for Biogeochemical Cycling and Hypoxia. Estuar. Coasts 44, 103–122. https://doi.org/10.1007/s12237-020-00763-8.

Nangir, D., Andredaki, M., Carnacina, I., 2025. Assessing the Feasibility of Satellite-Based Machine Learning for Turbidity Estimation in the Dynamic Mersey Estuary (Case Study: River Mersey, UK). Remote Sens. 17, 3617. https://doi.org/10.3390/rs17213617.

Raschka, Sebastian, Mirjalili, Vahid, 2019. Python machine learning : machine learning and deep learning with python, scikit-learn, and tensorflow 2. Packt Publishing, Limited.

Roberts, D.R., Bahn, V., Ciuti, S., Boyce, M.S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J.J., Schröder, B., Thuiller, W., Warton, D.I., Wintle, B.A., Hartig, F., Dormann, C.F., 2017. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography 40, 913–929. https://doi.org/10.1111/ecog.02881.

Robins, P.E., Skov, M.W., Lewis, M.J., Giménez, L., Davies, A.G., Malham, S.K., Neill, S.P., McDonald, J.E., Whitton, T.A., Jackson, S.E., Jago, C.F., 2016. Impact of climate change on UK estuaries: A review of past trends and potential projections. Estuar. Coast. Shelf Sci. 169, 119–135. https://doi.org/10.1016/j.ecss.2015.12.016.

Salama, Mhd.S., Spaias, L., Poser, K., Peters, S., Laanen, M., 2022. Validation of Sentinel-2 (MSI) and Sentinel-3 (OLCI) Water Quality Products in Turbid Estuaries Using Fixed Monitoring Stations. Frontiers in Remote Sensing 2. https://doi.org/10.3389/frsen.2021.808287.

Skarbøvik, E., Gyritia Madsen van't Veen, S., Lannergård, E.E., Wenng, H., Stutter, M., Bieroza, M., Atcheson, K., Jordan, P., Fölster, J., Mellander, P.-E., Kronvang, B., Marttila, H., Kaste, Ø., Lepistö, A., Kämäri, M., 2023. Comparing in situ turbidity sensor measurements as a proxy for suspended sediments in North-Western European streams. Catena (Amst). 225, 107006. https://doi.org/10.1016/j.catena.2023.107006.

Spieckermann, M., Gröngröft, A., Karrasch, M., Neumann, A., Eschenbach, A., 2022. Oxygen Consumption of Resuspended Sediments of the Upper Elbe Estuary: Process Identification and Prognosis. Aquat. Geochem. 28, 1–25. https://doi.org/10.1007/s10498-021-09401-6.

Strobl, C., Boulesteix, A.-L., Zeileis, A., Hothorn, T., 2007. Bias in random forest variable importance measures: Illustrations, sources and a solution. BMC Bioinformatics 8, 25. https://doi.org/10.1186/1471-2105-8-25.

Turner, A., Millward, G.E., 2002. Suspended Particles: Their Role in Estuarine Biogeochemical Cycles. Estuar. Coast. Shelf Sci. 55, 857–883. https://doi.org/10.1006/ecss.2002.1033.

Tyralis, H., Papacharalampous, G., Langousis, A., 2019. A Brief Review of Random Forests for Water Scientists and Practitioners and Their Recent History in Water Resources. Water (Basel). 11, 910. https://doi.org/10.3390/w11050910.

Vörösmarty, C.J., McIntyre, P.B., Gessner, M.O., Dudgeon, D., Prusevich, A., Green, P., Glidden, S., Bunn, S.E., Sullivan, C.A., Liermann, C.R., Davies, P.M., 2010. Erratum: Global threats to human water security and river biodiversity. Nature 468, 334. https://doi.org/10.1038/nature09549.

Winterwerp, J.C., 2002. On the flocculation and settling velocity of estuarine mud. Cont. Shelf Res. 22, 1339–1360. https://doi.org/10.1016/S0278-4343(02)00010-9.

Zhu, M., Wang, J., Yang, X., Zhang, Y., Zhang, L., Ren, H., Wu, B., Ye, L., 2022b. A review of the application of machine learning in water quality evaluation. Eco-Environment & Health 1, 107–116. https://doi.org/10.1016/j.eehl.2022.06.001.

Zhu, X., Guo, H., Huang, J.J., Tian, S., Xu, W., Mai, Y., 2022a. An ensemble machine learning model for water quality estimation in coastal area based on remote sensing imagery. J. Environ. Manag. 323, 116187. https://doi.org/10.1016/j.jenvman.2022.116187.