Continental Shelf Research 287 (2025) 105429

Contents lists available at ScienceDirect

Continental Shelf Research

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

Quantifying the relative contributions of forcings to the variability of estuarine surface suspended sediments using a machine learning framework

Juliana Tavora [a,f,*], Roy El Hourany [b], Elisa Helena Fernandes [c], Isabel Jalón-Rojas [d], Aldo Sotollichio [d], Mhd Suhyb Salama [a], Daphne van der Wal [a,e]

[a] Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, PO Box 217, 7500 AE, Enschede, the Netherlands
[b] Univ. Littoral Côte D'Opale, Univ. Lille, CNRS, IRD, UMR 8187, LOG, Laboratoire D'Océanologie et de Géosciences, F 62930, Wimereux, France
[c] Laboratório de Oceanografia Costeira e Estuarina, Instituto de Oceanografia, Universidade Federal Do Rio Grande (FURG), Rio Grande, Brazil
[d] Univ. Bordeaux, CNRS, Bordeaux INP, EPOC, UMR 5805, F-33600, Pessac, France
[e] Department of Estuarine and Delta Systems, NIOZ Royal Netherlands Institute for Sea Research, PO Box 140, 4400 AC, Yerseke, the Netherlands
[f] Ifremer, DYNECO-DHYSED, CS10070, 29280 Plouzane, France

A R T I C L E I N F O

Keywords:
turbidity$_1$
Surface sediment concentration$_2$
tide$_3$
River discharge$_4$
wind$_5$
Extreme events$_6$
Self organizing maps$_7$
Machine learning$_8$

A B S T R A C T

The influence of forcing mechanisms on the variability of suspended sediments in an estuary is, for the first time, synoptically quantified over prevailing ('normal') conditions and extreme events. This study investigates the complex and non-linear influence of tides, river discharge, and winds on the variability of suspended sediments in the macrotidal Gironde Estuary, France. Employing a machine learning-based framework, we integrated high-frequency field data, hourly numerical modeling outputs, and semi-daily satellite remote sensing to spatially quantify the relative contributions of forcing mechanisms. Our results reveal that tides are the primary driver of sediment variability (42.3–58.9%), followed by river discharge (21.2–34.7%) and wind (8.7–16.9%). Uncertainties range between 7% and 13.6%. In addition, the spatial variability of their contributions is consistent across numerical modeling and satellite remote sensing data, with differences not exceeding 10%. However, satellite data is limited by cloud cover and may miss extreme events. In contrast, hourly numerical modeling indicates tides are the dominant forcing mechanism under extreme events significantly affecting suspended sediment variability in the estuary. This study verifies the effectiveness of our machine learning approach against traditional Singular Spectral Analysis using field data. We demonstrate that machine learning techniques can effectively synthesize spatial distribution patterns of hydrodynamic and sedimentological variability, including the influence of winds. Our findings highlight not only the potential of satellite observations to analyze prevailing conditions despite data gaps but also that with hourly numerical modeling, the impact of forcings can be synoptically quantified under prevailing ('normal') conditions and extreme events.

1. Introduction

Estuaries, as dynamic transitional zones between riverine and marine environments, have an important role in buffering excess continental drainage and regulating the transport of nutrients (Wetz and Paerl, 2008), sediments, and other components like pathogens and pollutants (Robins et al., 2016) to the coastal sea. Much of the ability of estuaries to regulate transport between land and sea is also defined by the dynamic interplay of forcing mechanisms such as tides, river discharge, and winds (Meade, 1972; Dalrymple et al., 2012). These forcing mechanisms are critical in shaping the variability of sediment dynamics by influencing sediment transport, deposition, and resuspension patterns. For example, tides and river discharge influence the transport of sediments to the coastal zone. During ebbing tides, the exportation of sediments is favored, often resulting in turbid plumes, while flooding tides may modulate the entrainment of sediments in turbid maximum zones (TMZs; Doxaran et al., 2009). In addition, winds promote erosion and resuspension, especially in shallow regions (Constantin et al., 2018; Mulligan et al., 2019).

To effectively manage estuarine systems, it is critical to quantify the relative contributions of forcings mechanisms to suspended sediment variability in a range of hydrodynamic and meteorological settings.

* Corresponding author.
E-mail addresses: j.tavora@utwente.nl, juliana.tavora@ifremer.fr (J. Tavora).

https://doi.org/10.1016/j.csr.2025.105429
Received 26 April 2024; Received in revised form 9 December 2024; Accepted 12 February 2025
Available online 17 February 2025

Understanding these contributions is key to developing strategies aimed at adapting to (or mitigating) the impacts of human activities on the coast, such as dredging and port development, as well as climate change-driven impacts (Jalón-Rojas et al., 2017; e.g., extreme events) that are known for altering hydrodynamic circulation and by consequence sediment variability. Despite this need, efforts to measure and quantify the relative contributions of forcings to suspended sediments are still limited (French et al., 2008), especially in diverse hydrodynamic settings (e.g., 'normal' conditions and extreme events). Furthermore, the interaction of forcing mechanisms is not only central to sediment dynamics but also crucial for maintaining the ecological diversity and health of estuarine systems. Meteorological and oceanographic changes, ranging from prevailing ('normal') conditions to extreme events, are likely to alter the complexity of these interactions, thus posing significant challenges to the resilience of estuaries. Within the context of a global increase in the frequency and intensity of extreme events (Calvin et al., 2023), efforts to improve the understanding of the impacts of extreme events on coastal ecosystems are still evolving (e.g., Poppeschi et al., 2022; Tavora et al., 2023a) and efforts to develop adaptations are ongoing. In general, the primary effects of extreme events include but are not limited to heat waves, river floods, intensified wave conditions, and storm surges. Secondary impacts of extreme events include temporarily or permanently altered water quality, disruptions in the exportation of suspended sediment, particulate organic carbon, nutrients, and pollutants from land to the coastal ocean (e.g., Geyer et al., 2018; Wetz and Paerl, 2008; Wetz and Yoskowitz, 2013).

To understand and mitigate such disruptions to sediment dynamics in estuaries, either field data collection (from shipboard surveys or moored stations), observations from satellite sensors, or numerical model simulations are necessary. However, neither data source independently fully captures the variability driven by extreme events. For instance, satellite observations, while offering synoptic coverage, may be limited by clouds, tidal aliasing (Eleveld et al., 2014) or mismatched overpass timing during extreme events, and shipboard surveys may be unpractical due to unsafe meteorological and oceanographic conditions (Tavora et al., 2023a). While moored sampling stations provide high temporal data, their limited spatial coverage hinders their efficacy with discrete spatial sampling. Numerical models incorporating diverse temporal and spatial scales are used to overcome these challenges. However, refining their outcomes (for example, realistic flow of suspended sediment concentration (SSC) importation or accurate sediment class distribution), especially capturing extreme events, is challenging (Guinot and Gourbesville, 2003). By combining the different approaches, such as satellite, field observations, and numerical models, we maximize the extraction of information in the data sources and gain crucial insights into spatial and temporal dynamics. For instance, merging high-temporal-resolution data from moored stations and satellite observations, beyond just validation purposes, can provide essential information about high-frequency variability and spatial distribution of suspended sediments (Tavora et al., 2023a). Spectral analysis of long time series of high-frequency data from moored stations has shown the relative influence of environmental factors on the temporal fluctuations of suspended sediments within sub-hourly to daily scales (Jalón-Rojas et al., 2017; Schoellhamer, 2002). Satellite observations have been instrumental in tracking turbid river plumes, e.g., Tavora et al. (2023b), charting estuarine water quality, and investigating sediment resuspension driven by tidal fluctuation, winds, and river discharge within days to weeks (e.g., Tang et al., 2021). Numerical modeling contributes by providing hindcast or forecast scenarios at the water surface or by providing the hydrodynamical state of a system at depth.

However rich the synergy of all data sources, traditional data-driven analysis is often challenged due to noise, spatial or temporal gaps, or uneven sampling (Liang et al., 2023). These might hinder numerical modeling applications (which require input boundaries) and traditional statistical analysis. The latter has limitations assuming non-linear relationships and is frequently hampered by missing values and outliers (Liang et al., 2023). In contrast, the new generation of tools for data analysis using machine learning (ML) techniques provides extensive possibilities (Lary et al., 2016), particularly for problems where empirical knowledge is still incomplete but a comprehensive amount of data is available. As described by Lary et al. (2016), ML-based methods operate as "universal approximators" implying that these algorithms learn the relationship among variables from a set of training data without prior physical knowledge (Rubbens et al., 2023). That property has popularized the use of ML techniques in recent years for a variety of applications on water quality or its biogeochemical components. Examples include the works of Skákala et al. (2023) emulating hypoxia in the water column, Shin et al. (2020) predicting chlorophyll-*a*, Saccotelli et al. (2024) focused on salinity, and Nguyen et al. (2024) predicting suspended sediment concentration. The multiple possible applications and numerous ML approaches (e.g., neural networks, decision trees, random forests) – each technique with a trade-off among accuracy, interpretability, and data requirements – are further described in reviews of Lary et al. (2016), Rubbens et al. (2023), and Goldstein et al. (2019). Lary et al. (2016) address ML application in geoscience and remote sensing, Rubbens et al. (2023) describe ML applications and techniques for marine ecology, and Goldstein et al. (2019) discuss techniques applied to coastal sediment transport and morphodynamics. Here, we apply the unsupervised ML technique Self Organizing Maps (SOM) (Kohonen, 2013). The SOM is widely applied in meteorological and oceanographic studies (Liu and Weisberg, 2011) because it has been demonstrated to be an effective technique for pattern recognition and feature extraction, capable of capturing and reducing the complexity of non-linear patterns. The SOM's major advantage is the ability of reducing the dimensionality of the data while preserving the topological relationships among datapoints. This makes SOM useful for visualizing high-dimensional data in 2-dimensions, facilitating pattern recognition and clustering, while robustly dealing with noisy or incomplete data. However, SOM requires tuning of parameters like the number of learning neurons. SOM has, for example, facilitated the identification of phytoplankton groups (e.g., El Hourany et al., 2019; El Hourany et al., 2024; Yala et al., 2020), assessed uncertainties on sea level rise estimates (e.g., Camargo et al., 2023), and predicted oil spills (e.g., Mata et al., 2009).

In this study, we aim to quantify the interplay of meteorological and hydrodynamic forcings (i.e., river discharge, tides, and winds) over the variability of surface suspended sediments within the context of extreme events and prevailing ('normal') conditions. For this purpose, we propose a ML-based framework for estuaries and apply this to the Gironde Estuary, France. The proposed ML-based framework benefits from the inherent advantages of SOM handling high-dimensional data like hydrodynamic and sediment data. We feed the framework with numerical modeling and remote sensing data supplemented by field measurements. The framework allows the automatic assignment of different hydrodynamic and sediment distribution patterns within an estuary while attributing the influence of hydrodynamic forcings to the surface sediment variability at broad spatial scales, hereon referred to as synoptic scales.

2. Study site

The Gironde Estuary (Fig. 1), located on the Atlantic French coast, is the largest estuary in Western Europe. It exhibits a funnel-shaped system with increasing width (15 km at the mouth) towards the sea and about 75 km in length. The Gironde Estuary is a system extensively investigated under diverse hydrological conditions via numerical simulations (e.g., Diaz et al., 2020; Diaz et al., 2024; Jalón-Rojas et al., 2021), remote sensing methods (e.g., Constantin et al., 2017; Normandin et al., 2019), and high-resolution moored turbidity data (e.g., Etcheber et al., 2011; Jalón-Rojas et al., 2015; Jalón-Rojas et al., 2017).

The tidal regime in the estuary is characterized by high amplitudes (from 2.5m during neap tide to 5.5m during spring tide at Le Verdon).

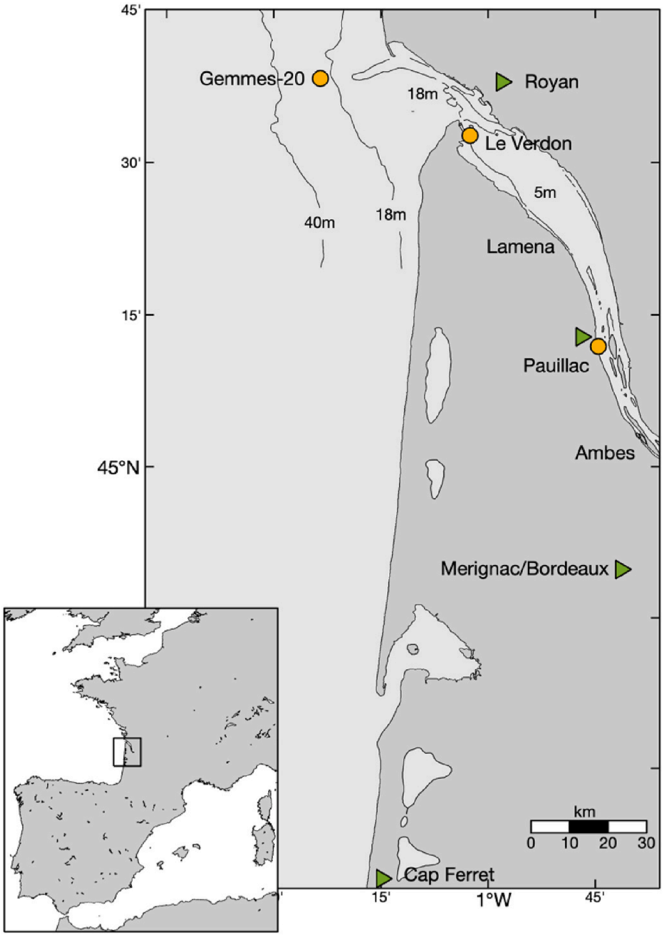


Fig. 1. Gironde Estuary, France. Green triangles represent the location of wind stations (Cap Ferret, Merignac/Bordeaux, Pauillac, and Royan). Dark yellow circles represent the location of turbidity data (from MAGEST program: Pauillac and Le Verdon, and from GEMMES buoy 20). Black contours represent the depth isobaths at 5m, 18m and 40m. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

The amplitude of the tides increases as one moves upstream, with a maximum range of 6.3m (Jalón-Rojas et al., 2018) being recorded at a distance of approximately 100–126 km from the estuary's mouth.

The Gironde Estuary receives freshwater inputs from two main rivers, the Garonne and the Dordogne rivers. The average freshwater discharge from these rivers is approximately 680 $m^3\ s^{-1}$, with reported peaks exceeding 4100 $m^3\ s^{-1}$ during winter. The discharge exhibits seasonal variations, with minimum values occurring between July and September (summer) and maximum values between December and February (winter). Winds from the northern sector prevail with average wind speeds between 2 and 4 $m\ s^{-1}$, rarely exceeding 10 $m\ s^{-1}$. The dynamics of winds at short time scales make their seasonal tendencies difficult to predict. In the inland stations, Pauillac and Merignac/Bordeaux, the mean wind speeds are approximately 4 $m\ s^{-1}$ and 3 $m\ s^{-1}$, respectively while in the Atlantic coast, at Cap Ferret station, wind speeds are up to three-folds stronger (refer to Section 3.5).

Previous studies have identified that tides and river discharge contribute up to 70% to the dynamics of sediments in the fluvial Gironde (Jalón-Rojas et al., 2017). As a result, the fluvial Gironde exhibits higher sediment concentrations during summer and lower concentrations during winter.

While tides and river discharge consistently influence sediment dynamics throughout most of the estuary (tides having a stronger influence), river discharge dominates during peak flow events in the fluvial Gironde (Jalón-Rojas et al., 2017). Winds are expected to have a significant role in the variability of sediments through wave resuspension, particularly in the shallow areas of the mouth and middle estuary (Constantin et al., 2017) such as intertidal mudflats, sandbanks, and shoals. Studies suggest the influence of wind events on sediment dynamics. Normandin et al. (2019) observed an inversely proportional relationship between winds and suspended sediments. Additionally, Lesueur et al. (2002) reported storm swells exceeding 15m, offshore of the mouth of the Gironde Estuary, highlighting the potential impact of storms on the estuary.

Regarding extreme events, on average at least one storm reaches the Gironde Estuary region every year (i.e., MeteoFrance (http://tempetes.meteo.fr/spip.php?article119; accessed on August 2023) provides information on storms), some of which caused widespread disruptions (e.g., storms Xynthia in 2010 (Kolen et al., 2013; Kolen et al., 2010), and Hercules in 2013–2014 (Castelle et al., 2015)). These storms, through intensified forcing mechanisms (such as winds) may influence sediment dynamics and turbidity levels in the estuary (Lesueur et al., 2002), or erosion of coastal areas (Castelle et al., 2015) and tidal flats, for example.

3. Material and methods

In this section we introduce the datasets used to compose the two main databases (based on hourly numerical model outputs and semi-daily satellite data) used for analysis of surface suspended sediment and the forcings related to their dynamics. We also introduce the ML techniques used to compose our framework for hydro-sedimentological regionalization. The approach handles a full time series of hydro-sedimentological data representing non-linear surface sediment responses to estuarine forcings along the Gironde Estuary.

3.1. Field data

Field data comprised time series of river discharge, wind speed, and turbidity. Fig. 1 illustrates the location of wind speed and turbidity (as a proxy for SSC) stations (see Supplementary Material A for summary of all field data). Hydrometric stations record daily discharges of the Dordogne River and the Garonne River (EAUFRANCE; http://www.hydro.eaufrance.fr/). Daily river discharge between 1959 and 2021 from each river was summed to achieve total daily amount of water flowing toward the coast.

Wind speed and direction were retrieved from four field stations (Royan, Pauillac, Merignac, and Cap Ferret) operated by the official service of meteorology and climatology in France (MétéoFrance; https://donneespubliques.meteofrance.fr/). Wind speed and wind direction were measured every hour at Royan (between years 1991 and 2022), Pauillac (2004–2020), and Merignac (2004–2020) stations and every 3h at Cap Ferret (1985–2022).

Field turbidity measurements were provided by MAGEST (MArel Gironde ESTuary; https://magest.oasu.u-bordeaux.fr) and GEMMES (Grasso et al., 2021). The MAGEST data were sampled at stations Le Verdon at the mouth of the Gironde estuary (between 2017 and 2021) and Pauillac in the middle estuary (52 km from the mouth, 2017–2018) and GEMMES data were sampled between 2016 and 2017 in front of Gironde Estuary mouth (buoy 20). The stations record turbidity at the water surface (<1m depth) every 10–30 min.

3.2. Remote sensing and derived turbidity estimates

A time series of all available, full resolution, level 1 data from Sentinel 3A, B (2016–2021) was processed using the ACOLITE processor (version 20221114; https://odnature.naturalsciences.be/remsem/software-and-data/acolite) from the Royal Belgium Institute of Natural Sciences (RBINS). The atmospheric correction default dark spectrum fitting scheme was applied to obtain water reflectance (R_w) and flags indicating atmospheric artifacts (e.g., non-water pixels, aerosol optical thickness out of range, or negative R_w). Every satellite scene was mapped to a cylindrical equidistant projection at a spatial resolution of 300m, and pixels attributed a flag indicating any atmospheric artifacts were removed.

Further, we carried out a match-up analysis between satellite overpass and field turbidity following the practices suggested in IOCCG (2019) and Tavora et al. (2023c). For match-ups we (i) opted for the minimal time difference between satellite overpass and field data acquisition, never surpassing 30min, and (ii) extracted the pixel centered at the location of the field station. Regarding the Pauillac station, the pixel centered at the location of the station was masked as land due to its proximity to the shore and the spatial resolution of the satellite sensor. Alternatively, the pixel coordinate was shifted to the nearest non-masked pixel (about 300m), a solution used in similar studies (e.g., Vanhellemont, 2019).

Satellite-derived turbidity was estimated by applying a Singular Vector Machine (SVM) regression between field turbidity data and R_w in seven satellite bands (620, 665, 709, 768, 779, 865, 884 nm). The use of multiple satellite bands is known for increasing the sensitivity to turbidity (e.g., Tavora et al., 2020) while avoiding saturation due to higher sediment concentrations commonly observed in the Gironde Estuary. Selection of such reflectance bands excluded the range 680–700 nm to avoid contamination of phytoplankton fluorescence in the rare cases of low turbidity. In addition, outliers represented by values below the 3rd percentile of the data distribution were removed, which corresponds to a threshold of 0.64 NTU. Refer to Supplementary Material B for metrics of satellite-derived turbidity.

3.3. Reanalysis wind data

Wind data was obtained from ECMWF-ERA5 (Hersbach et al., 2023) for the years 1959–2021 with 0.25^o spatial resolution. Wind data comprises both u and v components, 10m above surface with hourly frequency. To ensure that the dataset was not underestimating local wind speeds, we performed Cumulative Distribution Function (CDF) matching for bias correction (Singh et al., 2020; Reichle and Koster, 2004). ECMWF-ERA5 winds were compared for each wind component against each of the four field wind stations (i.e., Royan, Pauillac, Merignac/Bordeaux, and Cap Ferret). Refer to Supplementary Material C.

3.4. Hydrodynamical modeled data

The "CurviGironde" dataset (Diaz et al., 2023) was used. This data set consisted of output from the hydrodynamic model MARS-3D and sediment module MUSTANG, coupled to hindcast hydro-sediment dynamics between 2011 and 2021. Details of the model are described in Diaz et al. (2020). The modeled outputs were available in 10 equidistant sigma levels and a mesh grid ranging from 2 km offshore to 100m in the inner estuary. Diaz et al. (2020) validated the "CurviGironde" model output for hydrodynamic parameters (water level, tidal currents, wave height, salinity) with an ADCP and multiple tide gauging stations along the Gironde Estuary, and validated SSC with the MAGEST turbidity dataset at Pauillac station. Refer to Diaz et al. (2020) for validation metrics. The present study makes use of variables: depth, water level, tidal currents, and SSC between years 2016 and 2021, for comparability with satellite data. Tidal currents were further inspected for flooding and ebbing tidal directions, for which the data points relative to ebbing (flooding) tides were attributed a negative (positive) sign.

3.5. Detection of extreme events in the Gironde Estuary

An extreme event can refer to a value that exceeds a threshold determined from statistical methodologies, experimentally, or based on expert judgment (Stephenson, 2008). Here, we applied the peak-over-threshold to detect extreme events in the river discharge data. In the present study, we applied the 90th percentile threshold to detect extreme river discharge for data recorded between 1959 and 2021, as applied by Poppeschi et al. (2022) in the Bay of Brest and the Bay of Vilaine (France). The definition of windstorms in France; i.e., the absolute value of 25 m s^{-1} determined by MeteoFrance (http://tempetes.meteo.fr/Tempetes-cyclones-tornades-et-orages.html; accessed on August 2023), would lead to a complete misdetection of such events within the Gironde Estuary (refer to Supplementary Material D for maximum wind speeds within this study). Therefore, instead of windstorms, windbursts were further used. To determine windbursts, spatially integrated wind speed data between 1959 and 2021 (from bias-corrected ECMWF-ERA5) should surpass the 90th percentile of the time series.

3.6. The framework of hydro-sedimentological dynamics

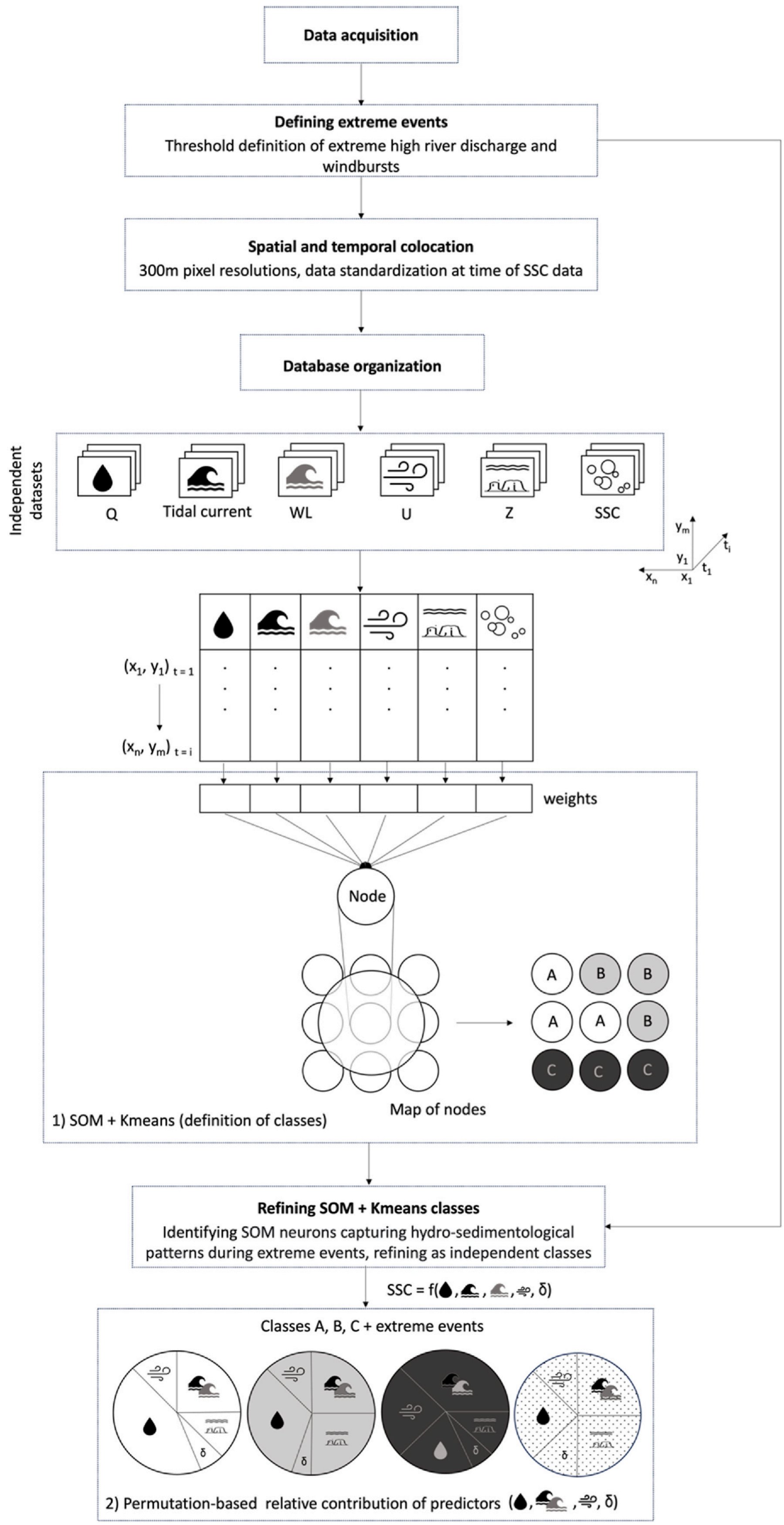
We propose here a framework (Fig. 2) to analyze the diverse hydro-sedimentological scenarios of estuaries in the Gironde Estuary. The framework and data analysis described here were carried out with MATLAB software (version 2023b). We first describe the database organization for training and verification (Section 3.6.1) and demonstrate the framework's ability to synthesize a database into hydro-sedimentological coherent regions while discriminating extreme events from prevailing conditions (Section 3.6.2). Finally, a permutation-based approach was used to determine the contribution of each forcing mechanism to sediment variability (Section 3.6.3). The process is carried out for model-based outputs and for satellite-based outputs.

3.6.1. Database organization and processing

The goal was to prepare variables (river discharge, tidal currents, water level, wind speed, depth, and a proxy for sediment dynamics – either SSC or turbidity) from independent data sources that were later used to provide base for the proposed analysis. These data were available under different temporal and spatial resolutions making the standardization step necessary. To achieve spatial comparability all variables were remapped onto an equidistant 300m resolution grid, using the spatial grid of satellite data as reference. For river discharge data, in addition to remapping, we performed spatial interpolation accounting for the time difference of 1 day between the location of gauging stations and the mouth of the estuary (Constantin et al., 2018).

Two datasets were further organized: Databases A and B. Database A consisted of a broad range of values and conditions for variables depth, forcings (river discharge, tides, winds) and model-based total surface sediment concentration (SSC). Database B consisted of the same variables as database A except for SSC, which was replaced by satellite-based turbidity. To achieve temporal comparability each database was aligned according to the most limiting data within its database. Database A was aligned at hourly temporal resolution (reaching over 50,000 time steps), and Database B was aligned with the timing of (partially) cloud-free satellite overpass (total of 216 satellite scenes). This standardization step, for each database individually, allowed for the integration of different data into a spatially uniform datacube (i.e., 3D arrays) organized by time, longitude, and latitude.

Each datacube was further reshaped in a database of 2D arrays such that each pixel (i.e., specific latitude, longitude and time) was considered a sampled point rearranged as a row vector and each column was relative to one of the six variables. These comprehensive 2D array databases consisted of, in order, hydrodynamical forcings along with depth and their respective surrogate of sediment dynamics. The two newly formed 2D array databases were then inspected for missing data and respective rows were removed. Finally, we applied log10-transformation to river discharge, SSC, and turbidity due to the data's broad range spanning three to four orders of magnitude.



(caption on next page)

Fig. 2. Sequence of approaches adopted for the proposed framework from data acquisition, pre-processing and application. Within the framework datacubes (dimensions $x_{1:n}$,$y_{1:m}$,$t_{1:i}$) of independent variables: river discharge (Q), tidal current, water level (WL), wind speed (U), depth (z), and surface sediment concentration (SSC) are reshaped into a 2D array. The 2D array was used as input to a SOM + Kmeans approach, and distinct regions (or classes) are determined. Further classes are refined for extreme events. Finally, with classes defined, the relative contribution of river discharge, tides (sum of contribution of tidal currents and water level), U, and uncertainty or noise (δ) is attributed to each recently determined class. Note that the number of classes (A, B, C + extreme events) depicted in this flowchart is a simplified representation and does not reflect the actual number of hydro-sedimentological classes identified in this study.

3.6.2. Hydro-sedimentological classification approach

Initially, a machine learning-based classification of hydro-sedimentological regimes in the Gironde Estuary was conducted using a two-step classification procedure. The first step involved using SOM (https://github.com/ilarinieminen/SOM-Toolbox/, accessed in August 2023) neural network used to reduce the complexity of features of the database in a 2D neural map while preserving the topology and capturing patterns within the neural map. At the end of the learning process, SOM positions similar units closely together while units representing different patterns are located further apart in the neural map. The SOM is an unsupervised classifier used for classification, noise reduction, and outlier detection. In the SOM analysis, due to the large amount of data, each database was subset with about 600,000 randomly selected data points. Further, variables in databases A and B were normalized to their respective variance to be similarly weighted and a SOM was trained. The SOM requires selecting an optimal number of neurons, which can be empirically determined in a trial-and-error approach (Marchese et al., 2022) as the value that provides the best representation of the training with minimized topographical (how well the topological map preserves the neighborhood relationships of the input data) and quantization errors (the difference between the input data and its closest assigned neuron). Minimizing both errors ensures that the SOM not only accurately represents the input data but also conserves the relationships between data points, leading to a more meaningful and interpretable topological map. We employed an approximate starting point for a number of neurons M, as applied by Elizondo et al. (2021) and Vesanto and Alhoniemi (2000), based on the total number of data points in the training database (n):

$$M = 5\sqrt{n} \qquad \text{Eq. 1}$$

By employing Eq. (1), we derived four initial reference values (i.e., M, M/2, 2M, and 4M) and adopted the reference value providing minimum topographical and quantization errors. In the second step, the Kmeans clusterization algorithm was used to reduce the number of clusters from the initial SOM, as suggested in Vesanto and Alhoniemi (2000). Using Kmeans after SOM is beneficial because it highlights hydro-sedimentological patterns in the reduced and organized space created by SOM. By preserving important patterns, SOM enables Kmeans to effectively reduce the number of clusters without sacrificing key relationships (El Hourany et al., 2019, 2024), addressing a common challenge in Kmeans cluster analysis. These clusters generated from SOM + Kmeans are hereon called hydro-sedimentological classes. By classifying data into hydro-sedimentological classes, we simplified the complexity of the data, grouping detailed, high-dimensional data into simpler, low-dimensional groups with similar characteristics in terms of water and sediment behavior. We further verified the detection of extreme events with the SOM + Kmeans by identifying SOM neurons that captured hydro-sedimentological patterns during extreme events with thresholds defined in Section 3.5. These were then refined and presented as independent classes.

Finally, statistical tests were conducted to assess whether the SOM + Kmeans hydro-sedimentological classes had distinct characteristics. First, a non-parametric, one-way analysis of variance by ranks, the Kruskal-Wallis H test, was conducted on each variable within the identified clusters, as in Marchese et al. (2022). This analysis was based on hydro-sedimentological attributes computed from each class's database. Complementarily, a post-hoc pairwise testing, the Dunn-Sidaks's multicomparison test, was applied to identify which classes differ significantly from the others, repeated to each variable (statistics are described in Supplementary Material E). Within the Dunn-Sidak's results, if at least one variable (either river discharge, tidal current, water level, wind speed, depth or suspended sediment) for a given class was statistically significant different from the other classes, then that class was different from the remaining classes.

3.6.3. Permutation-based relative importance of forcings to suspended sediments along estuary

To explain and estimate the relative importance of the predictors (river discharge, tides, and winds) to surface sediments, we applied a permutation-based importance method to a Random Forest (RF) algorithm using the MATLAB Statistics and Machine Learning toolbox (v23.2). This method, while computationally efficient for large datasets and of easy interpretability, involves randomly shuffling each predictor to evaluate its individual impact on surface suspended sediment prediction, measuring the resultant change in the RF model's performance. For each class in the Gironde Estuary, we trained a Random Forest algorithm using the option optimizable ensemble of trees with a default hyperparameter search range to predict the proxies of surface suspended sediment (SSC or turbidity), assuming a causal effect (refer to Data Statement to access trained RF algorithms; refer to Supplementary Material F for performance metrics of RF algorithms). Subsequently, we computed the mean relative importance of each predictor variable to establish their respective contributions. Lastly, each fitted RF algorithm was provided with a coefficient of determination (R^2) ranging between 0 and 1, meaning that the RF-algorithm explains 0–100% of the relationship between the proxy for surface sediment and predictor variables. Each R^2 was pondered in the relative contributions of predictor variables to account for stochastic signal of data as uncertainty and noise. Further, the uncertainties and noise were propagated to synoptical composites.

3.6.4. Intercomparison exercise for framework verification

Each database (A and B) consists of a source of information for suspended sediment dynamics independent from one another (i.e., model-based SSC and satellite-based turbidity), and each capturing different instants in the time series (random sub-set 1h intervals of model-based data versus the semi-daily sun-synchronous satellite data). For that purpose, we assess whether the framework applied to the different databases provides consistent relative contributions synoptically, independently of time intervals. The proposed framework was first applied to database A, which counts with a broader range of values for each variable. Then, the proposed hydro-sedimentological framework is independently applied to database B. For each database, the relative contribution of forcings per hydro-sedimentological class and synoptical relative contributions are compared. This comparison allows for assessing whether the machine learning-based framework provides consistent results, which is especially advantageous for data with gaps like satellite sensors.

Relative contribution results with the framework are compared with the Singular Spectral Analysis (SSA) described by Jalón-Rojas et al. (2016a) and Jalón-Rojas et al. (2017). The SSA decomposes a time series into a set of principal components of nearly periodic oscillations with the contribution of the variance of each principal component determined by its eigenvalue. While the SSA is shown to detect the periodic oscillations such as observed in tides and river discharge, other environmental forcings, such as winds, are more likely detected as noise rather than explained variance, as observed in Jalón-Rojas et al. (2017) for the

Gironde Estuary and Loire Estuary (FRA). We applied SSA to Le Verdon (2017–2021) following the methodology applied and described in Jalón-Rojas et al. (2017) and compared it with the framework's estimated relative contribution centered at the position of the station. The SSA was not applied to Pauillac and GEMMES stations due to the short time frame (Jalón-Rojas et al., 2016a), preventing a more accurate comparison between methods for relative contribution.

4. Results

4.1. Characteristics and patterns of identified hydro-sedimentological classes

The SOM + Kmeans algorithm distinguished seven hydro-sedimentological classes that exhibit the most significant partition of the database (database A). Of these, a few classes capture characteristics of extreme river discharge events, (potential) windbursts, or capture extreme river discharge and potential windbursts simultaneously. The hydro-sedimentological classes identified capturing extreme events were then refined and split into classes specific for extreme river discharge, windbursts, and their simultaneous occurrence. Refining these classes to separately capture extreme events allows the detection (and discrimination) of classes that are affected by extreme events and classes that are not (prevailing conditions) while also providing an overview of associated environmental patterns. Overall, a total of 10 hydro-sedimentological classes (i.e., 7 classes of prevailing conditions and 3 classes of extreme events) were established summarizing the complex and highly dynamic system of the Gironde Estuary. The established number of classes is supported by both the Kruskal-Wallis H (p-value <0.05) and Dunn-Sidak's multicomparison test (Q-value <0.05). The Kruskal-Wallis and Dunn-Sidaks's test revealed that for a given class at least one variable (river discharge, tidal current, water level, wind speed, depth, or SSC/turbidity) was statistically significant different from the remainder classes (refer to results in Supplementary Material E).

Fig. 3 depicts the 10 hydro-sedimentological classes grouped in prevailing conditions (C1-C7; Fig. 4a–f) and extreme events (C8-C10; Fig. 3g–l). Each class is generally characterized by different surface sediment concentration ranges with an interchangeable combination of low, intermediate, or high values for the remaining variables. For example, while classes C3 and C4 yield similarities in most variables, they differ in terms of water level and river discharge with C3 representing lower water levels and more intense river discharge than C4.

4.2. Quantification of forcings to surface suspended sediment variability

The dynamic nature of estuarine sediments here represented by SSC (database A), is attributed to a variability of forcings including river discharge, tides, and winds for each class. According to the results obtained from the framework, each and all forcings contribute to predicted SSC (Fig. 4). Among classes of prevailing conditions, tides (42.3%–58.9%) and river discharge (21.2%–34.7%) are most frequently observed as the dominant contributors to SSC variability, with higher relative contribution compared to winds (8.7%–16.9%) and uncertainty or noise (7%–13.6%). Classes representing extreme events yield similar relative contributions for forcings as classes representing prevailing conditions.

4.3. Synoptical relative contributions to surface suspended sediment variability

Once the relative contribution to surface sediment variability is identified per class, we can identify the relative contribution of forcings at any timestep, and the overall synoptical relative contributions of the entire timeseries. Fig. 5 depicts the overall synoptical contributions, using database A, for which river discharge (Fig. 5a) and tides (Fig. 5b), as well as winds (Fig. 5c) and uncertainties (Fig. 5d) show different ranges of contribution to suspended sediment variability. Wind and uncertainties are of secondary importance compared to river discharge and tides. It is also clear that the spatial variation in the relative contributions of the variables is low. However, the land-to-sea continuum displays some short-range variability, with a smooth increasing trend on the tidal contribution (from a mean of 47.4% $\pm$ 0.8% in the inner estuary to a mean of 51.5% $\pm$ 1.8% in the outer estuary). River discharge shows a small consistent gradient in the continuum, with a slightly stronger contribution in the inner estuary (mean of 29.3% $\pm$ 0.7%) but decreasing seawards (mean of 26.7% $\pm$ 1.1%). Winds demonstrate a slightly stronger contribution to SSC variability in the shallower areas, including the inner estuary and sand banks or shoals on the outer estuary (mean of 15.1% $\pm$ 0.3%) as opposed to a mean of 13.5% $\pm$ 0.6% in the deeper areas of the outer estuary. Uncertainties or noise are also roughly consistent spatially (about 8.3% $\pm$ 0.4%).

4.3.1. Comparison with Singular Spectral Analysis

Further, we compare results from the proposed hydro-sedimentological framework with database A to those of the established SSA approach with field turbidity data. Table 1 shows the relative contributions of predicting factors to sediment variability with both methods. Results are relatively consistent between methods, particularly for river discharge and uncertainty (differences $<$4.5%), while tides yield about 15% difference in relative contribution to SSC variability. Finally, the hydro-sedimentological framework provides estimates for the influence of winds on the variability of sediments in suspension. At the same time, for SSA, the impact of winds is unquantified, due to its stochastic nature, but likely contained within uncertainty estimates.

4.4. Database intercomparison verification

We compare the relative contribution results obtained from the hydro-sedimentological framework applied to database A and database B (see Supplementary Material G-1 for the relative contribution of classes with database B). Results achieved with each database are consistent, with less than 10% difference for either parameter (Fig. 6). Overall, results obtained from database A yield slightly lower relative contribution for river discharge and wind speeds. Tides and uncertainties, on the other hand, yielded slightly higher relative contributions to SSC variability. In general, the database intercomparison exercise demonstrates that the hydro-sedimentological framework yields similar results for the overall relative contribution of forcings and uncertainty, regardless of the source of suspended sediment data (hourly numerical model or satellite-based).

The largest differences between databases A and B are related to detecting relative contributions under extreme events. Contrarily to database A, with database B, the forcings representing extreme events (e.g., river discharge for extreme river discharge or winds for potential windbursts) consisted of a short range of values detected during (partially) cloud-free satellite overpass in the occurrence of extreme events (i.e., 5 scenes under extreme river discharge, 34 scenes of potential windburst, and 1 scene under simultaneous extreme events). This limited range of river discharge or wind speed values likely constrains the permutation-based approach due to the low number of degrees of freedom on retrieving relative contributions under extreme events (Supplementary Material G −2,3).

4.5. Changes detected under extreme events in the Gironde Estuary

Although the analysis addressed extreme events, it reveals that tides are the primary driver of sediment variability. The influence of tides is greater than the effect of extreme events (river discharge and windbursts, as shown in Fig. 4). Additionally, the overall spatial relative contribution of forcings during extreme events reveals that tidal contributions are roughly consistently 15% higher than under prevailing

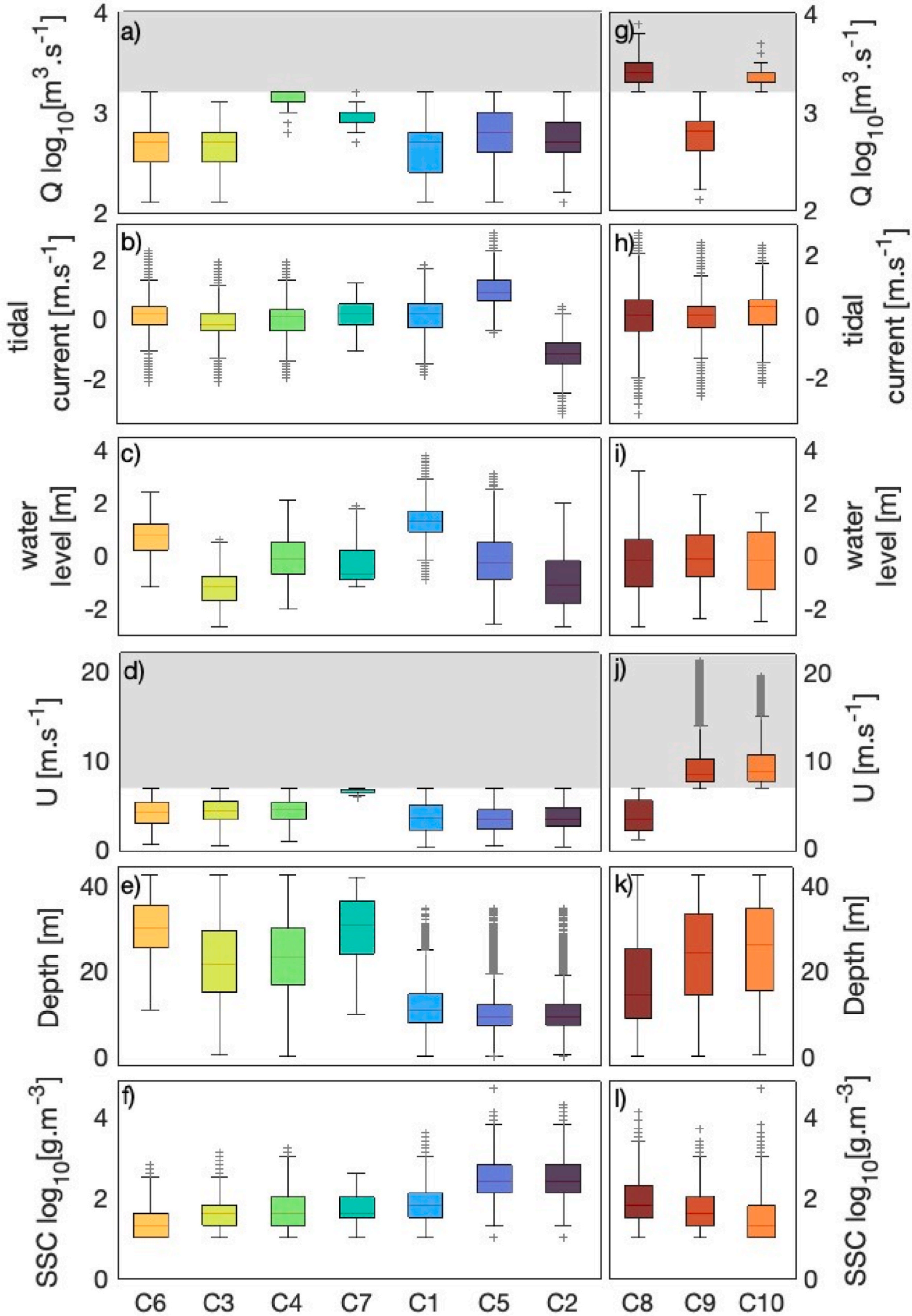


Fig. 3. Classes of coherent hydrodynamical and sedimentological patterns of database A (x-axis), for river discharge (Q $\log_{10}$ transformed; a,g), tidal current (b,h), water level (c,i), wind speed (U; d,j), depth (e,k) and model output of surface sediment concentration (SSC $\log_{10}$ transformed; f,l). Left panels (a–f) represent hydro-sedimentological classes under prevailing conditions, while right panels (g–i) represent classes under extreme event conditions (respectively simultaneous extreme river discharge, potential windburst, and simultaneous extreme river discharge and potential windburst). The shaded upper portion of (a,g) and (d,j) respectively represent the threshold defining occurrence of extreme river discharge ($Q_{\log10} > 3.24$ m^3 s^{-1}) and potential windburst events ($U > 6.82$ m s^{-1}) in the Gironde Estuary. Classes of prevailing conditions are ordered based on increasing SSC.

Fig. 4. Class-specific relative contribution (%) of predicting factors (river discharge, tide, wind and uncertainty or noise) for modeled surface SSC (database A). Classes C1-C7 depict relative contribution of forcing mechanisms under prevailing conditions (ordered from low (C6) to high (C2) log SSC), classes C8-C10 depict relative contribution of extreme events (extreme high river discharge, potential windbursts, and simultaneous extremes, respectively).

conditions whereas river discharge and winds have a lower relative contribution to sediment variability during extreme events.

A further investigation of parameters representing tides (i.e., tidal currents and water level) shows that water levels under extremes are above those observed under prevailing conditions (Fig. 7a–c), with differences reaching up to 0.5m. The spatial patterns observed for tidal currents (Fig. 7d–f) are more complex showing a clear influence of bathymetric features. Among the diverse patterns, mean tidal currents under extreme river discharge (Fig. 7d) are stronger than under prevailing conditions. Potential windbursts, however, result in lower mean tidal currents, especially in the deeper regions of the estuary like the navigation channel. The contrast between prevailing and extreme conditions is stronger during simultaneous extremes, splitting the inner estuary into three parts: (i) the region between Le Verdon and Lamena with stronger currents under simultaneous extremes, (ii) the region between Lamena and Pauillac with weaker currents under extremes, and (iii) a region upward Pauillac station with a smooth gradient of increasing tidal current speed under simultaneous extremes. The outer estuary yielded similar patterns between windbursts and simultaneous extremes, differing on the magnitude of tidal currents while spatial patterns observed for extreme river discharge are roughly the opposite.

The dynamics of forcing mechanisms reflect the variability of suspended sediments. Fig. 8 depicts the spatial differences in SSC by type of extreme event: river discharge (Fig. 8a), potential windbursts (Fig. 8b), and simultaneous extreme events (Fig. 8c). The influence of forcing mechanisms is observed along the entire estuarine domain, yet each type of extreme event leads to a distinct SSC distribution. Extreme river discharge, for instance, yields intensified SSC in the middle estuary and low SSC within the most upstream and downstream regions. Potential windbursts, compared to prevailing conditions, yield higher SSC throughout the estuary. SSC under simultaneous extremes reveals a mix of patterns observed under extreme river discharge and potential windbursts with a well-marked (and more extensive) region of higher SSC and plume-like features in the outer estuary.

5. Discussion

5.1. Perspectives on the use of machine learning to address the influence of forcings to estuarine suspended sediment dynamics

In this study, we have employed turbidity field data to validate the accuracy of satellite-based turbidity estimates (field data was also used for verification of the “CurviGironde” numerical model outputs by Diaz et al., 2020). We then compared results from the proposed ML framework applied to satellite-based information and to the model output “CurviGironde” to ensure consistency.

Our study used an unsupervised neural network ML (the Self Organizing Maps – SOM + Kmeans) combined with an explanatory approach to analyze discontinuous satellite remote sensing data and consistent hourly numerical model outputs. This novel method not only matched the results from traditional approaches, i.e., SSA, but provided advantages over these traditional methods, such as (1) spatially resolved estimates of the influence of forcings on suspended sediment variability (consistent from both numerical model outputs and satellite remote sensing data), and (2) accounted (and estimated) for uncertainties while (3) providing the influence of winds on sediment variability. However, despite the advantages, the framework applied to satellite remote sensing of Sentinel 3 (OLCI) alone is limited in capturing estuarine dynamics under extreme events due to limited temporal data coverage. Instead, by leveraging a high temporal and spatial data source (i.e., numerical model output), we added a fourth advantage of the proposed ML framework: (4) isolating the influence of forcings and the resulting sediment response during prevailing conditions compared to extreme events. In comparison to the traditional statistical approach, SSA, which is mostly limited by gaps and length of time series as discussed in Jalón-Rojas et al. (2016a), and the fact that SSA cannot distinguish the role of winds, the proposed machine learning-based framework, has distinct advantages. The advantages observed with the proposed ML framework over traditional methods are attributed to the key skills of ML, especially regarding SOM. These are largely attributed to capturing

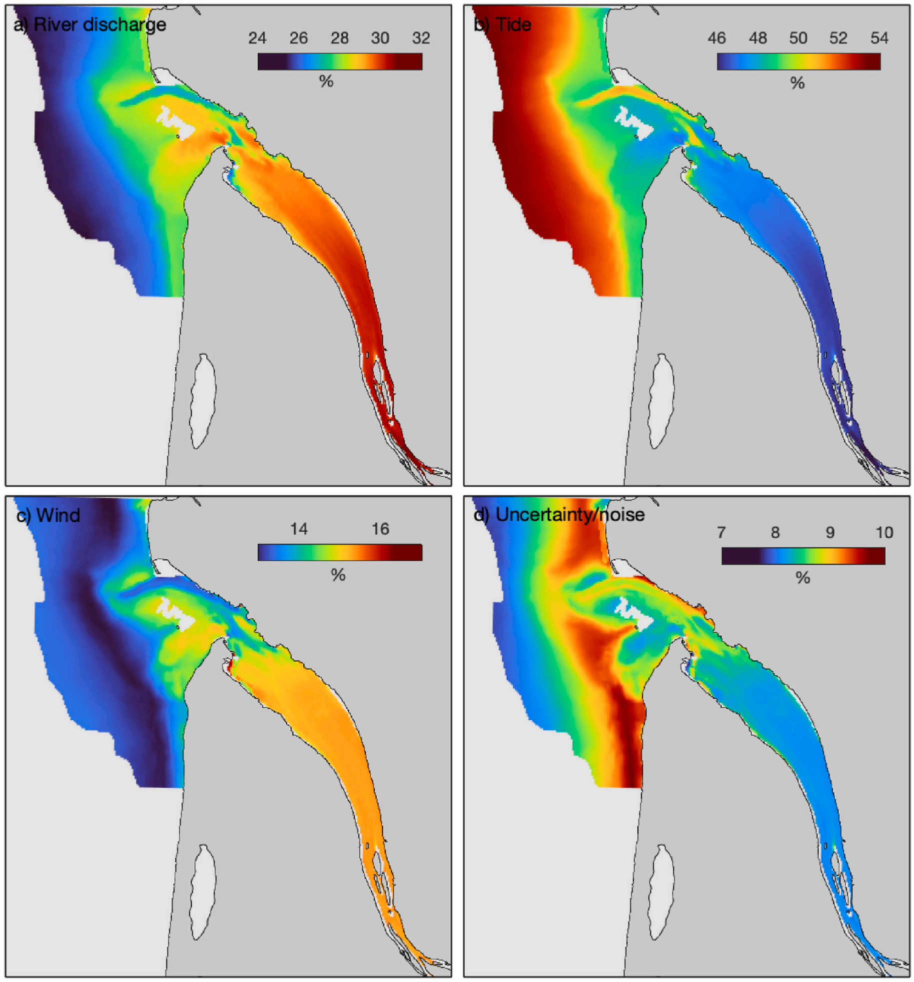


Fig. 5. Synoptical mean contribution of predicting factors for modeled surface SSC (database A): (a) river discharge, (b) tides, (c) winds and (d) uncertainties/noise to variability in surface suspended sediment concentrations. Attention to the different ranges of percentage on each subplot.

Table 1
Relative contribution of forcings and uncertainty (or noise) to suspended sediment variability from the hydro-sedimentological framework (from database A) and the Singular Spectral Analysis (SSA) approach (from field data).

Station	Predictor	Hydro-sedimentological framework [%]	SSA [%]
Le Verdon	river discharge	28.4	23.9
	tide	48.3	63.5
	wind	14.4	–
	uncertainty	8.9	12.6

hydro-sedimentological patterns despite the uneven sampled data inputs, gaps in the input data, and the topology preservation of SOM that reduces the complexity of the hydro-sedimentological data while keeping key relationships.

The Gironde Estuary was a perfect study site, counting with long-term and high-frequency data at fixed stations (especially from the MAGEST program), which aided in the verification of the proposed framework but also with a high-frequency (validated) numerical model output (Diaz et al., 2023) providing means to address extreme events (and their influence on sediment dynamics). While the Gironde Estuary was an ideal laboratory site to develop and test the hydro-sedimentological ML framework, the proposed framework may be employed in other estuarine environments using satellite remote sensing data. This versatility for spatial quantification of the influence of forcing mechanisms is exemplified by the comparable results obtained with hourly numerical modeling (database A) and satellite remote sensing (database B), paving the way for comprehensive environmental studies.

5.2. Interplay of forcing mechanisms to sediment variability in the Gironde Estuary

It is not new that on a macrotidal estuary like the Gironde Estuary, tides play a major role in sediment variability. Previous studies with field sampled data (e.g., Castaing and Allen, 1981; Jalón-Rojas et al., 2017; Jalón-Rojas et al., 2016b), satellite remote sensing (e.g., Constantin et al., 2018; Normandin et al., 2019) or numerical modeling (e.g., Diaz et al., 2020), have attributed much of the sediment dynamics firstly to tides and secondly to river discharge – some with qualitative or with quantitative approaches but most of them using traditional statistical analysis or numerical modeling. While backed up by the literature, the present study aims to complement the state-of-art by providing a machine learning-based framework to quantify the complex role of major forcings to sediment variability, spatially and temporally, while also quantitatively attributing the role of winds. Results suggest small spatial ranges of relative contributions. Additionally, compared against the Singular Spectral Analysis (SSA) for the contribution of river discharge and tides, the framework suggests agreement with the role of the major forcings and that winds play a small part in the variability of surface sediments. Except for the navigation channel within the mouth of the estuary (Fig. 1; depth >18m), results show that the influence of tides decreases up estuary, with a decreasing trend consistent with

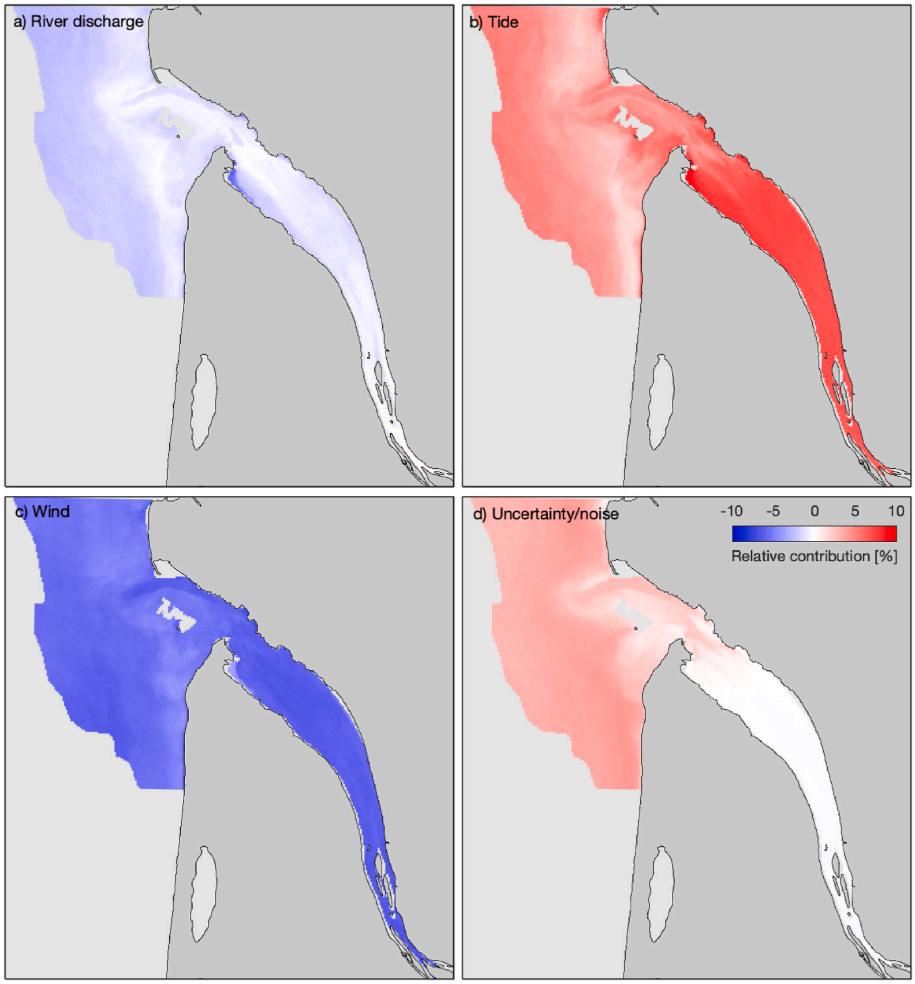


Fig. 6. Spatial difference between overall relative contribution for variability of surface sediment dynamics estimated from database A and database B. Red (blue) shades indicate higher overall relative contributions estimated from database A (B). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

previous studies in the Gironde Estuary and Loire Estuary (FRA; Jalón-Rojas et al., 2017). Right on the navigation channel, at the estuary's mouth, the tidal contribution is higher than its shallower surroundings, likely due to stronger tidal velocity gradients. A stronger tidal velocity gradient in the navigation channel was also reported by Diaz et al. (2020) and Ross et al. (2019). With opposite trends, river discharge yielded increasing contributions towards the upper estuary. An increasing trend was also observed with SSA using field turbidity data between 2004 and 2015 in the fluvial section of the Gironde Estuary (Jalón-Rojas et al., 2017). The intercomparison exercise shows agreement between results applying the framework to numerical model-based data (database A) and satellite-based data (database B) for prevailing conditions. The most significant difference (<10%) is observed in the shallow tidal mudflat between stations Le Verdon and Lamena for river discharge contribution. These results suggest that satellite remote sensing, despite spatial and temporal gaps, applied to the proposed framework can provide spatial estimates of relative contribution of forcings to sediment variability.

As river discharge and tides are the major forcing mechanisms in the Gironde Estuary, the importance of winds and their effects on sediment dynamics have been less studied. SSA, with high-temporal resolution field data, cannot distinguish the role of winds from the stochastic signal (uncertainty or noise). Through satellite remote sensing, Constantin et al. (2018) attempted to retrieve the effects of winds on the turbid coastal plumes of the Gironde Estuary, but results were inconclusive, while Normandin et al. (2019) observed a weak to moderate negative relationship between winds and a principal component of suspended sediment concentration. Our proposed hydro-sedimentological framework provides yet another advantage by spatially resolving the role of wind on the variability of surface sediments. Our results show a low influence of winds along estuary with lower influence in the deeper regions at the mouth of the estuary (<14%), coinciding with the navigation channel. In addition, the offshore region of the Gironde Estuary (depth >40m) experiences lower influence of winds compared to the shallower regions of the inner estuary (depth <18m), where we observed the highest influence of winds (>17%). Similar patterns of stronger influence of winds in shallow areas including sandbanks, shoals, and intertidal mudflats were reported by Normandin et al. (2019). Doxaran et al. (2009) also report that winds are likely more important in these shallow areas contributing to erosion and consistent sediment resuspension.

5.3. Suspended sediments in the context of extreme events

During extreme events, the Gironde Estuary experiences higher SSC along the estuarine continuum. However, the spatial patterns of SSC differ among types of extreme events. Of all three types of extreme events, the spatial pattern under extreme river discharge depicts SSC below prevailing conditions except in the middle estuary (between Lamena and Pauillac) indicating the presence of a TMZ. The lower SSC observed in the estuary's lower- and upper-most regions suggests that sediments are remobilized seawards, coinciding with stronger ebbing

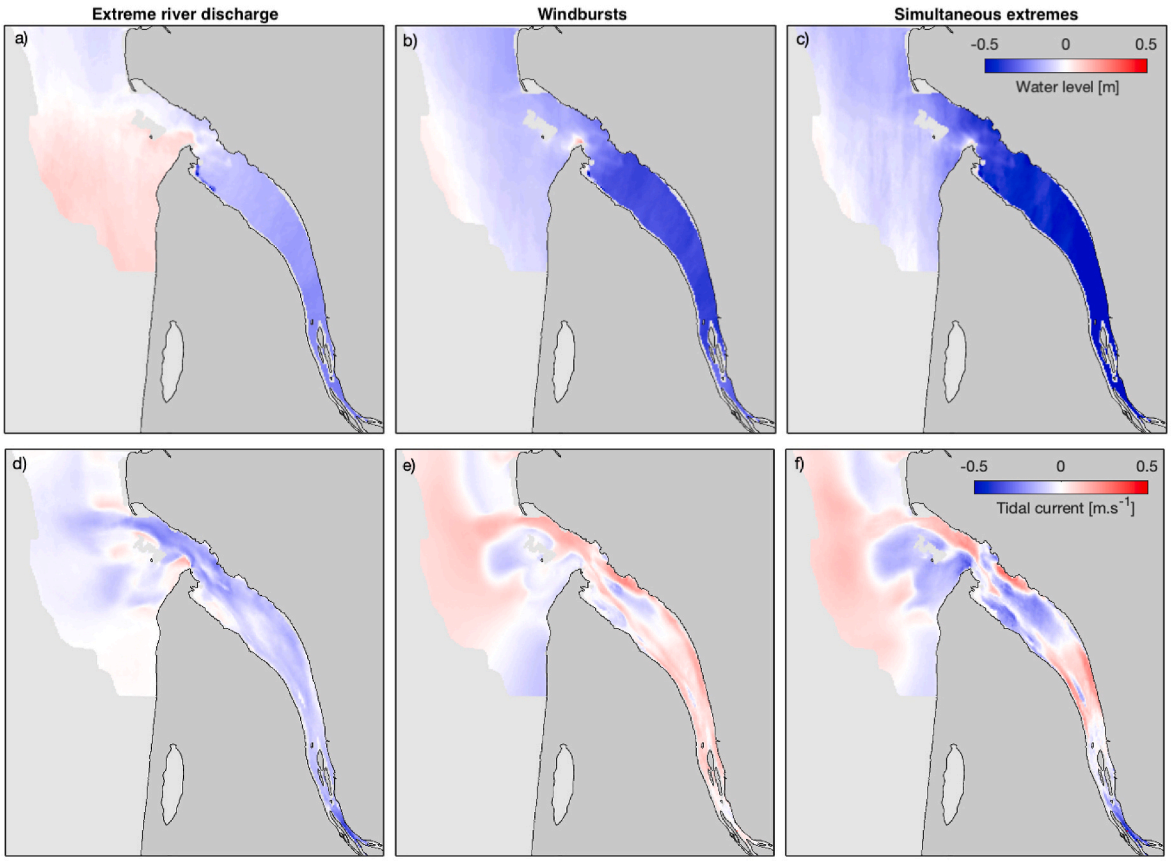


Fig. 7. Spatial difference between prevailing and extreme conditions of water level (a–c) and tidal current (d–f). Regions in blue (red) depict lowest (highest) water level or tidal currents under prevailing conditions. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

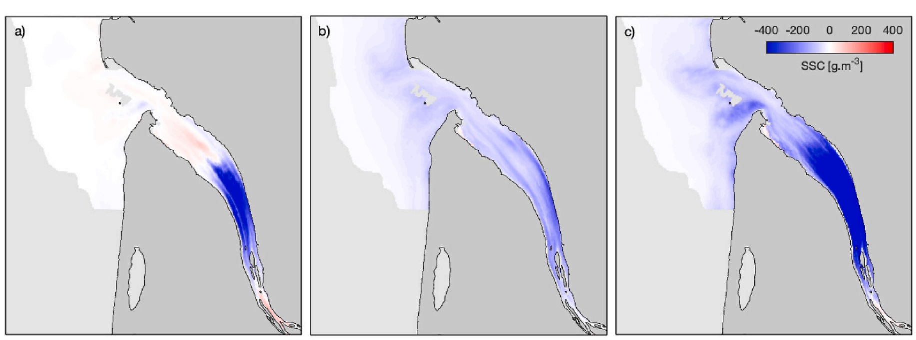


Fig. 8. Spatial difference between SSC under prevailing conditions and SSC under (a) extreme river discharge, (b) windbursts, and (c) simultaneous occurrence of extreme river discharge and windbursts. Regions in blue (red) depict lowest (highest) SSC under prevailing conditions. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

tidal currents due to more intense river discharge and lower friction due to higher water levels. In the section of the estuary between Le Verdon and Lamena, sediments are transported seaward, as suggested by Castaing and Allen (1981) and Doxaran et al. (2009). Similarly, in the narrow, most fluvial part of the estuary, the region between stations Pauillac and Ambes (but likely extending further upstream as reported by Jalón-Rojas et al., 2015) indicates overall seaward (downstream) migration of the TMZ from the most fluvial to the middle Gironde Estuary. This migration of the TMZ has been vastly explained and discussed (e.g., Jalón-Rojas et al., 2015; Normandin et al., 2019; Sottolichio and Castaing, 1999).

Contrary to the pattern observed within extreme river discharge, SSC values under windbursts are observed to be larger than those under prevailing conditions throughout the estuary (depth <18m). This difference in SSC suggests that surface sediments remain suspended in shallow areas due to wind-action, either remobilizing sediments (or preventing sediment settling) or weakening tidal currents. SSC is the highest in the navigation channel (depths >18m), and patterns coincide with mostly weaker tidal currents under windbursts. Among all three types of extremes investigated, the simultaneous occurrence of potential windbursts and extreme river discharge yielded the highest overall SSC. In addition, SSC patterns under simultaneous extremes reveal similarities with patterns observed both under extreme river discharge and windbursts. During simultaneous extreme events, an extensive and more concentrated TMZ is found. In the outer estuary, where the turbid plume is often observed (Constantin et al., 2018), SSC was detected above prevailing conditions in both potential windburst and simultaneous extremes. Those regions of relatively higher SSC coincide with stronger outflowing tidal currents.

Tidal currents present a complex spatial pattern of alternate high/low speeds. Overall, strong tides were observed for extreme river discharge when the water level is slightly higher. The condition may be due to a reduced friction effect of tides in higher water levels and, consequently, stronger tides. This was similarly described in San Francisco Bay by Holleman and Stacey (2014). The weakest tidal currents are observed during windbursts, likely due to winds opposing or reducing the ebbing currents (and trapping waters within the estuary). Still, spatial patterns of tidal currents are relatively similar on both potential windbursts and simultaneous extremes, especially in the outer estuary. While tidal currents show variability within extremes, water levels are similar among types of extreme events: higher (lower) water levels under extreme events in the inner (outer) estuary. The difference in water level relies mainly on where, in the longitudinal extension of the estuary, water levels are above or below prevailing conditions.

5.4. Limitations and future work

The study addressed the interplay and influence of forcing mechanisms (river discharge, tides, and winds) on sediment variability but was limited at the surface. Using numerical model outputs like the "CurviGironde", the interplay of forcings on sediment variability can also be estimated at depth, providing additional insights into sediment dynamics for the full water column. However, this was beyond the scope of this study. We anticipate, however, that under well-mixed water column conditions, patterns identified at the surface reflect the patterns at depth, but over stratified water conditions the proposed framework likely offers novel perspectives on bottom sediment dynamics. Being the Gironde Estuary well-mixed in spring tides and stratified in neap tides (Allen et al., 1980), we can expect that the patterns and estimated influence of forcings observed during spring tides or very shallow regions reflect those at depth, and that during neap tides the framework may provide new insights also at depth.

Looking ahead, the time frame applied to the study, from 2016 to 2021, limits the scope for climatological analyses, which require longer time frames (ideally >30 years). However, consistent application of the framework to extended time frames may indicate an evolution of meteorological and oceanographic scenarios and impacts over sediment dynamics. Further, the use of a geostationary satellite sensor, or alternatively, the combination of various satellite sensors, i.e., Landsat 8 (Operational Land Imager), Landsat 9 (Operational Land Imager), Sentinel 2 A/B (Multispectral Instrument), and Sentinel 3 A/B (Ocean and Land Color Instrument), applied to the proposed framework may provide comprehensive information for estimating the influence of forcings to the variability sediments including more coverage of extreme events, although not tested here.

6. Conclusions

This study contributes to understanding the spatial influence of forcing mechanisms between suspended sediments under prevailing ('normal') conditions and under extreme events of river discharge, potential windbursts, and their simultaneous occurrence with potential for application in estuaries worldwide. From this study, we conclude that (1) the proposed machine learning framework matches the performance of traditional methods, such as Singular Spectral Analysis at a single station, while also providing spatially-resolved estimates of relative contributions of forcings under extreme events; (2) the proposed machine learning framework has the additional advantage of estimating the contributions of winds to surface sediment variability and providing uncertainty estimates; (3) a total of 10 classes suffice to characterize the hydro-sedimentological variability within the Gironde Estuary; (4) tide is the main forcing mechanism controlling sediment variability in the Gironde Estuary; (5) the relative contribution of forcing mechanisms on the variability of suspended sediments exhibits a small spatial variability; (6) the relative role of tides increases during extreme events; and (7) winds have the lowest influence on surface sediment variability.

CRediT authorship contribution statement

Juliana Tavora: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Roy El Hourany:** Writing – review & editing, Methodology, Formal analysis, Conceptualization. **Elisa Helena Fernandes:** Writing – review & editing, Data curation. **Isabel Jalón-Rojas:** Writing – review & editing. **Aldo Sotollichio:** Writing – review & editing, Resources, Data curation. **Mhd Suhyb Salama:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Investigation, Conceptualization. **Daphne van der Wal:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Investigation, Conceptualization.

Data availability statement

The datasets and RF algorithms generated for this study will be found in the DANS (Data Archiving and Networked Services) repository. The original field data (turbidity, winds, and river discharge) are under license, thus protected. CurviGironde data by Diaz et al. (2023) is under license CC-BY-NC-SA 4.0 (for non-commercial use and share alike https://creativecommons.org/licenses/by-sa/4.0/). The results on this manuscript contain modified Copernicus Climate Change Service information. Neither the European Commission nor ECMWF is responsible for any use that may be made of the Copernicus information or data it contains. Remaining original data are open source. Finally, the Hydro-sedimentological framework will be available as a package of MATLAB functions under https://github.com/julianatavora/HYDROSED_framework.

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.

Acknowledgments

This work used data provided by many sources. Turbidity data come from the GEMMES (Grasso et al., 2021) and the MAGEST network (https://magest.oasu.u-bordeaux.fr/), maintained and by Sabine Schmidt (CNRS). Sentinel 3 comes from Scihub COPERNICUS, HydroEAU (https://hydro.eaufrance.fr/) provided river discharge data, winds come from MeteoFrance (https://meteofrance.com/) and ECMWF-ERA5 (Hersbach et al., 2023) downloaded from Copernicus Climate Change Service. Authors are grateful to SUNSET Project - South Brazilian Shelf sediment transport: sources and consequences (CAPES/COFECUB, Process 88887.192855/2018-00). Elisa Helena Fernandes is a CNPq research fellow (PQ2 No 304684/2022-8). Finally, the authors also thank Mariana Belgiu for her reassuring perspectives on using machine learning.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.csr.2025.105429.

Data availability

Data will be made available on request.

References

Allen, G.P., Salomon, J.C., Bassoullet, P., Du Penhoat, Y., de Grandpré, C., 1980. Effects of tides on mixing and suspended sediment transport in macrotidal estuaries. Sediment. Geol. 26 (1–3), 69–90. https://doi.org/10.1016/0037-0738(80)90006-8.

Calvin, K., Dasgupta, D., Krinner, G., Mukherji, A., Thorne, P.W., Trisos, C., et al., 2023. In: Arias, P., Bustamante, M., Elgizouli, I., Flato, G., Howden, M., Méndez-Vallejo, C., et al. (Eds.), IPCC, 2023: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)], IPCC. https://doi.org/10.59327/IPCC/AR6-9789291691647. Geneva, Switzerland.

Camargo, C.M.L., Riva, R.E.M., Hermans, T.H.J., Schütt, E.M., Marcos, M., Hernandez-Carrasco, I., Slangen, A.B.A., 2023. Regionalizing the sea-level budget with machine learning techniques. Ocean Sci. 19 (1), 17–41. https://doi.org/10.5194/os-19-17-2023.

Castaing, P., Allen, G.P., 1981. Mechanisms controlling seaward escape of suspended sediment from the Gironde: a macrotidal estuary in France. Mar. Geol. 40 (1–2), 101–118. https://doi.org/10.1016/0025-3227(81)90045-1.

Castelle, B., Marieu, V., Bujan, S., Splinter, K.D., Robinet, A., Sénéchal, N., Ferreira, S., 2015. Impact of the winter 2013-2014 series of severe Western Europe storms on a double-barred sandy coast: beach and dune erosion and megacusp embayments. Geomorphology 238, 135–148. https://doi.org/10.1016/j.geomorph.2015.03.006.

Constantin, S., Constantinescu, Ştefan, Doxaran, D., 2017. Long-term analysis of turbidity patterns in Danube Delta coastal area based on MODIS satellite data. J. Mar. Syst. 170, 10–21. https://doi.org/10.1016/j.jmarsys.2017.01.016.

Constantin, S., Doxaran, D., Derkacheva, A., Novoa, S., Lavigne, H., 2018. Multi-temporal dynamics of suspended particulate matter in a macro-tidal river Plume (the Gironde) as observed by satellite data. Estuar. Coast Shelf Sci. 202 (January 2020), 172–184. https://doi.org/10.1016/j.ecss.2018.01.004.

Dalrymple, R.W., Mackay, D.A., Ichaso, A.A., Choi, K.S., 2012. Processes, morphodynamics, and facies of tide-dominated estuaries. In: Davis Jr., R., Dalrymple, R. (Eds.), Principles of Tidal Sedimentology. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-0123-6_5.

Diaz, M., Grasso, F., Le Hir, P., Sottolichio, A., Caillaud, M., Thouvenin, B., 2020. Modeling mud and sand transfers between a macrotidal estuary and the continental shelf: influence of the sediment transport parameterization. J. Geophys. Res.: Oceans 125 (4), 1–37. https://doi.org/10.1029/2019JC015643.

Diaz, Melanie, Grasso, F., Le Hir, P., Sottolichio, A., Caillaud, M., 2024. Investigating sediment dynamics on a continental shelf mud patch under the influence of a macrotidal estuary: a numerical modeling analysis. Cont. Shelf Res. 282. https://doi.org/10.1016/j.csr.2024.105334.

Diaz, Mélanie, Grasso, F., Caillaud, M., 2023. CurviGironde hindcast. IFREMER. https://doi.org/10.12770/44ac4d72-c606-42ba-bf22-89e6520e0894.

Doxaran, D., Froidefond, J.M., Castaing, P., Babin, M., 2009. Dynamics of the turbidity maximum zone in a macrotidal estuary (the Gironde, France): observations from field and MODIS satellite data. Estuar. Coast Shelf Sci. 81 (3), 321–332. https://doi.org/10.1016/j.ecss.2008.11.013.

Eleveld, M.A., van der Wal, D., Van Kessel, T., 2014. Estuarine suspended particulate matter concentrations from sun-synchronous satellite remote sensing: tidal and meteorological effects and biases. Remote Sensing of Environment 143, 204–215. https://doi.org/10.1016/j.rse.2013.12.019.

Elizondo, U.H., Righetti, D., Benedetti, F., Vogt, M., 2021. Biome partitioning of the global ocean based on phytoplankton biogeography. Prog. Oceanogr. 194 (January 2020), 102530. https://doi.org/10.1016/j.pocean.2021.102530.

Etcheber, H., Schmidt, S., Sottolichio, A., Maneux, E., Chabaux, G., Escalier, J.M., et al., 2011. Monitoring water quality in estuarine environments: lessons from the MAGEST monitoring program in the Gironde fluvial-estuarine system. Hydrol. Earth Syst. Sci. 15 (3), 831–840. https://doi.org/10.5194/hess-15-831-2011.

French, J., Burningham, H., Benson, T., 2008. Tidal and meteorological forcing of suspended sediment flux in a muddy mesotidal estuary. Estuaries Coasts 31 (5), 843–859. https://doi.org/10.1007/s12237-008-9072-5.

Geyer, N., Huettel, M., Wetz, M., 2018. Biogeochemistry of a river-dominated estuary influenced by drought and storms. Estuaries Coasts 41 (7), 2009–2023. https://doi.org/10.1007/s12237-018-0411-x.

Goldstein, E.B., Coco, G., Plant, N.G., 2019. A review of machine learning applications to coastal sediment transport and morphodynamics. Earth Sci. Rev. 194 (4), 97–108. https://doi.org/10.1016/j.earscirev.2019.04.022.

Grasso, F., Bocher, A., Jacquet, M., Le Berre, D., Lecornu, F., 2021. Gironde estuary mouth MEasurement stations (GEMMES). SEANOE. https://doi.org/10.17882/78968.

Guinot, V., Gourbesville, P., 2003. Calibration of physically based models: back to basics? J. Hydroinform. 5 (4), 233–244. https://doi.org/10.2166/hydro.2003.0020.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., et al., 2023. ERA5 hourly data on single levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). https://doi.org/10.24381/cds.adbb2d47.

Holleman, R.C., Stacey, M.T., 2014. Coupling of sea level rise, tidal amplification, and inundation. J. Phys. Oceanogr. 44 (5), 1439–1455. https://doi.org/10.1175/JPO-D-13-0214.1.

El Hourany, R., Abboud-Abi Saab, M., Faour, G., Aumont, O., Crépon, M., Thiria, S., 2019. Estimation of secondary phytoplankton pigments from satellite observations using self-organizing maps (SOMs). J. Geophys. Res.: Oceans 124 (2), 1357–1378. https://doi.org/10.1029/2018JC014450.

El Hourany, R., Pierella Karlusich, J., Zinger, L., Loisel, H., Levy, M., Bowler, C., 2024. Linking satellites to genes with machine learning to estimate phytoplankton community structure from space. Ocean Sci. 20 (1), 217–239. https://doi.org/10.5194/os-20-217-2024.

IOCCG, 2019. Uncertainties in ocean colour remote sensing. In: Mélin, F. (Ed.), Report Ser). Dartmouth, Canada: International Ocean Color Coordinating Group.

Jalón-Rojas, I., Schmidt, S., Sottolichio, A., 2015. Turbidity in the fluvial Gironde Estuary (southwest France) based on 10-year continuous monitoring: sensitivity to hydrological conditions. Hydrol. Earth Syst. Sci. 19 (6), 2805–2819. https://doi.org/10.5194/hess-19-2805-2015.

Jalón-Rojas, I., Sottolichio, A., Hanquiez, V., Fort, A., Schmidt, S., 2018. To what extent multidecadal changes in morphology and fluvial discharge impact tide in a convergent (turbid) tidal river. J. Geophys. Res.: Oceans 123 (5), 3241–3258. https://doi.org/10.1002/2017JC013466.

Jalón-Rojas, I., Dijkstra, Y.M., Schuttelaars, H.M., Brouwer, R.L., Schmidt, S., Sottolichio, A., 2021. Multidecadal evolution of the turbidity maximum zone in a macrotidal river under climate and anthropogenic pressures. J. Geophys. Res.: Oceans 126 (5), 1–20. https://doi.org/10.1029/2020JC016273.

Jalón-Rojas, Isabel, Schmidt, S., Sottolichio, A., 2016a. Evaluation of spectral methods for high-frequency multiannual time series in coastal transitional waters: advantages of combined analyses. Limnol Oceanogr. Methods 14 (6), 381–396. https://doi.org/10.1002/lom3.10097.

Jalón-Rojas, Isabel, Schmidt, S., Sottolichio, A., Bertier, C., 2016b. Tracking the turbidity maximum zone in the Loire Estuary (France) based on a long-term, high-resolution and high-frequency monitoring network. Cont. Shelf Res. 117, 1–11. https://doi.org/10.1016/j.csr.2016.01.017.

Jalón-Rojas, Isabel, Schmidt, S., Sottolichio, A., 2017. Comparison of environmental forcings affecting suspended sediments variability in two macrotidal, highly-turbid estuaries. Estuar. Coast Shelf Sci. 198, 529–541. https://doi.org/10.1016/j.ecss.2017.02.017.

Kohonen, T., 2013. Essentials of the self-organizing map. Neural Netw. 37, 52–65. https://doi.org/10.1016/j.neunet.2012.09.018.

Kolen, B., Slomp, R., Jonkman, S.N., 2013. The impacts of storm Xynthia February 27-28, 2010 in France: lessons for flood risk management. Journal of Flood Risk Management 6. https://doi.org/10.1111/jfr3.12011.

Kolen, Bas Slomp, R., Balen, W. wan, Terpstra, T., Bottema, M., Nieuwenhuis, S., 2010. Learning from French experiences with storm Xynthia: damages after a flood. Rijksoverheid. ISBN 978-90-77051-77-1.

Lary, D.J., Alavi, A.H., Gandomi, A.H., Walker, A.L., 2016. Machine learning in geosciences and remote sensing. Geosci. Front. 7 (1), 3–10. https://doi.org/10.1016/j.gsf.2015.07.003.

Lesueur, P., Tastet, J.P., Weber, O., 2002. Origin and morphosedimentary evolution of fine-grained modern continental shelf deposits: the Gironde mud fields (Bay of Biscay, France). Sedimentology 49 (6), 1299–1320. https://doi.org/10.1046/j.1365-3091.2002.00498.x.

Liang, W., Liu, T., Wang, Y., Jiao, J.J., Gan, J., He, D., 2023. Spatiotemporal-aware machine learning approaches for dissolved oxygen prediction in coastal waters. Sci. Total Environ. 905 (September), 167138. https://doi.org/10.1016/j.scitotenv.2023.167138.

Liu, Y., Weisberg, R.H., 2011. A review of self-organizing map applications in meteorology and oceanography. In: Self Organizing Maps - Applications and Novel Algorithm Design. InTech. https://doi.org/10.5772/13146.

Marchese, C., Hunt, B.P.V., Giannini, F., Ehrler, M., Costa, M., 2022. Bioregionalization of the coastal and open oceans of British Columbia and Southeast Alaska based on Sentinel-3A satellite-derived phytoplankton seasonality. Front. Mar. Sci. 9 (September), 1–22. https://doi.org/10.3389/fmars.2022.968470.

Mata, A., Corchado, E., Baruque, B., 2009. Solving the oil spill problem using a combination of CBR and a summarization of SOM ensembles. International Symposium on Distributed Computing and Artificial Intelligence 2008 (DCAI 2008) 50, 658–662. https://doi.org/10.1007/978-3-540-85863-8_78. Berlin, Heidelberg: Springer Berlin Heidelberg.

Meade, R.H., 1972. Transport and deposition of sediments. Estuaries 91–120. https://doi.org/10.1130/MEM133-p91.

Mulligan, R.P., Smith, P.C., Tao, J., Hill, P.S., 2019. Wind-wave and tidally driven sediment resuspension in a macrotidal basin. Estuaries Coasts 42 (3), 641–654. https://doi.org/10.1007/s12237-018-00511-z.

Nguyen, S.Q., Nguyen, L.C., Ngo-Duc, T., Ouillon, S., 2024. Applying a machine learning-based method for the prediction of suspended sediment concentration in the Red river basin. Model. Earth Syst. Environ. 10, 2675–2692. https://doi.org/10.1007/s40808-023-01915-y.

Normandin, C., Lubac, B., Sottolichio, A., Frappart, F., Ygorra, B., Marieu, V., 2019. Analysis of suspended sediment variability in a large highly turbid estuary using a 5-year-long remotely sensed data archive at high resolution. J. Geophys. Res.: Oceans 124 (11). https://doi.org/10.1029/2019JC015417, 2019JC015417.

Poppeschi, C., Charria, G., Daniel, A., Verney, R., Rimmelin, P., Retho, M., et al., 2022. Interannual Variability of the Initiation of the Phytoplankton Growing Period in Two French Coastal Ecosystems. https://doi.org/10.5194/bg-2022-86 (April), 1–29.

Reichle, R.H., Koster, R.D., 2004. Bias reduction in short records of satellite soil moisture. Geophys. Res. Lett. 31, L19501. https://doi.org/10.1029/2004GL020938, 2004.

Robins, P.E., Skov, M.W., Lewis, M.J., Giménez, L., Davies, A.G., Malham, S.K., et al., 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.

Ross, L., Huguenard, K., Sottolichio, A., 2019. Intratidal and fortnightly variability of vertical mixing in a macrotidal estuary: the Gironde. J. Geophys. Res.: Oceans 124 (4), 2641–2659. https://doi.org/10.1029/2018JC014456.

Rubbens, P., Brodie, S., Cordier, T., Destro Barcellos, D., Devos, P., Fernandes-Salvador, J.A., et al., 2023. Machine learning in marine ecology: an overview of techniques and applications. ICES (Int. Counc. Explor. Sea) J. Mar. Sci. 80 (7), 1829–1853. https://doi.org/10.1093/icesjms/fsad100.

Saccotelli, L., Verri, G., De Lorenzis, A., Cherubini, C., Caccioppoli, R., Coppini, G., Maglietta, R., 2024. Enhancing estuary salinity prediction: a Machine Learning and Deep Learning based approach. Applied Computing and Geosciences. https://doi.org/10.1016/j.acags.2024.100173.

Schoellhamer, D.H., 2002. Variability of suspended-sediment concentration at tidal to annual time scales in San Francisco Bay, USA. Cont. Shelf Res. 22 (11–13), 1857–1866. https://doi.org/10.1016/S0278-4343(02)00042-0.

Shin, Y., Kim, T., Hong, S., Lee, S., Lee, E., Hong, S., Lee, C., Kim, T., Park, M.S., Park, J., et al., 2020. Prediction of chlorophyll-*a* concentrations in the nakdong river using machine learning methods. Water 12, 1822. https://doi.org/10.3390/w12061822.

Singh, A., Gaurav, K., Meena, G.K., Kumar, S., 2020. Estimation of soil moisture applying modified dubois model to sentinel-1; A regional study from Central India. Remote Sens 12, 2266. https://doi.org/10.3390/rs12142266.

Skákala, J., Awty-Carroll, K., Menon, P.P., Wang, K., Lessin, G., 2023. Future digital twins: emulating a highly complex marine biogeochemical model with machine learning to predict hypoxia. Front. Mar. Sci. 10, 1058837. https://doi.org/10.3389/fmars.2023.1058837.

Sottolichio, A., Castaing, P., 1999. A synthesis on seasonal dynamics of highly-concentrated structures in the Gironde estuary. Comptes Rendus Acad. Sci. - Ser. IIA Earth Planet. Sci. 329 (11), 795–800. https://doi.org/10.1016/S1251-8050(00)88634-6.

Stephenson, D.B., 2008. Definition, diagnosis, and origin of extreme weather and climate events. Climate Extremes and Society 9780521870, 11–23. https://doi.org/10.1017/CBO9780511535840.003.

Tang, R., Shen, F., Ge, J., Yang, S., Gao, W., 2021. Investigating typhoon impact on SSC through hourly satellite and real-time field observations: a case study of the Yangtze Estuary. Cont. Shelf Res. 224 (April), 104475. https://doi.org/10.1016/j.csr.2021.104475.

Tavora, J., Boss, E., Doxaran, D., Hill, P., 2020. An algorithm to estimate suspended particulate matter concentrations and associated uncertainties from remote sensing reflectance in coastal environments. Remote Sens. 12 (13), 1–24. https://doi.org/10.3390/rs12132172.

Tavora, J., Salama, Mhd S., Penning de Vries, M., Mannaerts, C.M., van der Wal, D., 2023a. Detecting the effects of extreme events on estuarine suspended particulate matter using satellite remote sensing (scheldt estuary): challenges and opportunities. Remote Sens. 15 (3), 670. https://doi.org/10.3390/rs15030670.

Tavora, J., Gonçalves, G.A., Fernandes, E.H., Salama, Mhd S., van der Wal, D., 2023b. Detecting turbid plumes from satellite remote sensing: state-of-art thresholds and the novel PLUMES algorithm. Front. Mar. Sci. 10 (July), 1–18. https://doi.org/10.3389/fmars.2023.1215327.

Tavora, J., Jiang, B., Kiffney, T., Bourdin, G., Gray, P.C., Carvalho, L.S., et al., 2023c. Recipes for the derivation of water quality parameters using the high-spatial-resolution data from sensors on board sentinel-2A, sentinel-2B, landsat-5, landsat-7, landsat-8, and landsat-9 satellites. Journal of Remote Sensing 3 (Table 1), 1–18. https://doi.org/10.34133/remotesensing.0049.

Vanhellemont, Q., 2019. Adaptation of the dark spectrum fitting atmospheric correction for aquatic applications of the Landsat and Sentinel-2 archives. Remote Sensing of Environment 225 (March), 175–192. https://doi.org/10.1016/j.rse.2019.03.010.

Vesanto, J., Alhoniemi, E., 2000. Clustering of the self-organizing map. IEEE Trans. Neural Network. 11 (3), 586–600. https://doi.org/10.1109/72.846731.

Wetz, M.S., Paerl, H.W., 2008. Estuarine phytoplankton responses to hurricanes and tropical storms with different characteristics (trajectory, rainfall, winds). Estuaries Coasts 31 (2), 419–429. https://doi.org/10.1007/s12237-008-9034-y.

Wetz, M.S., Yoskowitz, D.W., 2013. An "extreme" future for estuaries? Effects of extreme climatic events on estuarine water quality and ecology. Mar. Pollut. Bull. 69 (1–2), 7–18. https://doi.org/10.1016/j.marpolbul.2013.01.020.

Yala, K., Niang, N., Brajard, J., Mejia, C., Ouattara, M., El Hourany, R., et al., 2020. Estimation of phytoplankton pigments from ocean-color satellite observations in the Senegalo-Mauritanian region by using an advanced neural classifier. Ocean Sci. 16 (2), 513–533. https://doi.org/10.5194/os-16-513-2020.