Sediment texture and geochemistry as predictors of sub-depositional environment in a modern estuary using machine learning: A framework for investigating clay-coated sand grains

T.E.Nichols[a], R.H.Worden[a]∗, J.E.Houghton[a], R.A.Duller[a], J.Griffiths[ab], J.E.P.Utley[a]

a: Diagenesis Research Group, Department of Earth, Ocean, and Ecological Sciences, School of Environmental Sciences, University of Liverpool, 4 Brownlow Street, Liverpool L69 3GP, UK

b: National Nuclear Laboratory, 5th Floor, Chadwick House, Birchwood Park, Risley, Warrington WA3 6AE, UK

∗Corresponding author. rworden@liverpool.ac.uk.

Abstract

Sedimentary cores from the Ravenglass Estuary lack some of the sedimentary structures which can be seen in other estuarine sands due to their unconsolidated nature, making it difficult to meaningfully interpret depositional environments using standard sedimentological facies analysis. Here we explore how sediment texture, obtained from laser particle size analysis, and geochemistry, obtained from portable X-ray fluorescence, can be used independently, or in combination, to automatically classify sub-depositional environment and estuarine zone in a modern estuary. We have adapted an established Extreme Gradient Boosting workflow to select the most informative geochemical elements to be included in a training set to automatically classify sub-depositional environment at the surface of the Ravenglass Estuary, NW England, UK. Models that are trained exclusively on textural data significantly outperform those that use geochemical data when classifying sub-depositional environment but are comparable when classifying estuarine zone. However, the combination of textural and geochemical data in training sets improves model performance in all but one class when compared to separate textural and geochemical models. We have applied surface-calibrated combined textural and geochemical models to classify palaeo sub-depositional environment in three cores obtained from a tidal flat in the Ravenglass Estuary which that are interpreted to record initial outer estuary deposition which transitioned to an inner estuary setting dominated by deposition from suspension. The subsurface classifications provide a framework to investigate the occurrence of reservoir quality-preserving detrital grain coats which vary in abundance as a function of sub-depositional environment at the surface. A review of literature suggests that the mixed flat sub-depositional environment is an ideal target for optimum clay grain coat occurrence, and we show that this sub-depositional environment is reliably identified by all models. We also discuss the implications of the modelling, how it compares to other machine learning approaches to understand reservoir quality in ancient systems, and how the workflow may be adapted for application to reservoir core. This study demonstrates value in utilising textural and geochemical data in conjunction with machine learning methods to help reveal the environmental evolution of marginal marine sands.

Keywords: Marginal marine sands, Machine learning,Clay grain coats, Sub-depositional environment, Sediment texture, Sediment geochemistry

Data availability

Data to support this work for model creation and figure production is available in the open access repository here: https://doi.org/10.5281/zenodo.8419160.

1. Introduction

The observed heterogeneity in sediments across marginal marine systems arises because of variations in type and intensity of hydrodynamic processes, such as tidal, wave, and fluvial processes (Dalrymple et al., 1992; Boyd et al., 2006; Ainsworth et al., 2011). The composition (texture and geochemistry) of material supplied to a sedimentary system is controlled by many extrinsic factors including (but not limited to) hinterland geology, transport distance, and chemical weathering intensity (McLennan, 1993; Garzanti et al., 2009; Caracciolo, 2020). Sediment texture and geochemistry have been shown to be useful in determining these extrinsic factors in ancient systems, such as the use of element ratios to determine tectonic setting (Roser and Korsch, 1986), and relationships between grain size in fluvial systems to subsidence and sediment supply (Duller et al., 2010). However, in this study we will focus on how texture and geochemistry reflect internal heterogeneities of a marginal marine sedimentary system, where the distribution of different grain size fractions and minerals is controlled by intrinsic hydrodynamic sorting (Lambiase, 1980; Worden et al., 2020). At the surface, these heterogeneities produce a variety of landforms, sedimentary structures, and habitats for organisms (Dalrymple and Zaitlin, 1994; Virolle et al., 2020). Areas of an estuary surface which share these physical characteristics, and texture and geochemical properties can be used to define 'sub-depositional environments' (e.g. tidal inlet, mud flats, or foreshore) within a larger sedimentary environment (e.g., an estuary or delta) (Dalrymple and Zaitlin, 1994; Heap et al., 2004; Griffiths et al., 2019a).

When studying modern sedimentary systems at the surface, the identification of sub-depositional environment is straightforward as the features that define these environments can be observed unobstructed (e.g., sedimentary structures, flora, fauna, geomorphology, and geographic position). Here we define a palaeo-sub-depositional environment as a subsurface (>2 cm depth) expression of the equivalent surface sub-depositional environment. When looking into the subsurface of a modern environment, it is expected that some of the features that define sub-depositional environments are preserved in sediment cores expressed, for example, as ichnofabrics and sedimentary structures, enabling environmental and/or facies interpretation based on these features (Fenies and Tastet, 1998; Dalrymple and Choi, 2007; Virolle et al., 2019b; Virolle et al., 2020). However, sediment cores obtained from the Ravenglass Estuary in northwest England did not exhibit these expected structures as expected, particularly when sand-dominated material was recovered. The poor preservation of sedimentary structures and ichnofabrics, exacerbated by the sediment's poorly consolidated nature, made identifying palaeo-sub-depositional environment challenging. Therefore, a reliable method to determine palaeo-sub-depositional environment in sediment core, that is not reliant on good preservation of structures, was necessary to understand the sedimentary succession in the Ravenglass Estuary cores. Here we used texture (i.e., grain size characteristics) and geochemistry as universal properties of sediments to discriminate sub-depositional environment at the surface to create machine learning classification models (Simon et al., 2021; Muhammed et al., 2022; Houghton et al., 2023). The machine learning classification model can then be used to classify subsurface intervals, providing a likely environment of deposition even in structureless sand.

Machine learning is the use of computational methods to identify patterns in, and to classify, data without being explicitly programmed. Supervised machine learning algorithms train models to classify samples with a known label/class (e.g., sub-depositional environment) using associated data as predictors. In the case of sedimentary systems, the effects of hydrodynamic sorting can result in statistically significant differences in texture and geochemistry between sub-depositional environments of marginal marine sediments, a facet of the system that can be exploited by machine learning algorithms (Simon et al., 2021; Muhammed et al., 2022). The bulk geochemistry and texture of sediments are generally correlated due to the tendency of clay minerals and hydroxides to be most abundant as fine-grained material, concentrating elements such as Al and Fe, whereas coarse-grained material is commonly enriched in Si, dominant in quartz and feldspars (Nesbitt and Young, 1996; Guo et al., 2018). Therefore, obtaining textural and geochemical data from surface sedimentary environments facilitates model building for sub-depositional environment classification and prediction of subsurface environments. Whilest surface sub-depositional environment is not directly translatable into ancient buried systems, it is analogous to facies schemes which are common in core studies (Churchill et al., 2017; Griffiths et al., 2021) and can provide insight into the three-dimensional distribution of deposits. Geochemical logging tools have been used to classify downhole lithology and mineralogy using multivariate techniques; however, here, we apply a machine learning method to classify subsurface core intervals with genetic 'surface sub-depositional environment' labels (Herron, 1988; Hertzog et al., 1989; Freedman et al., 2015). Understanding the compositional characteristics of sediments (particularly sandstones) and the distribution of sub-depositional environment/facies is important in many real-world applications as they can be a significant control on porosity and permeability (i.e., reservoir quality) in terms of primary depositional character and influences on diagenetic activity (Beard and Weyl, 1973; Chan, 1985; Worden and Burley, 2003).

Upon deep burial (>2 km; >70 °C), sandstones typically develop diagenetic quartz cement which nucleates on the surface of detrital quartz grains and grows to occlude pore space (Walderhaug, 1996; Ajdukiewicz and Lander, 2010; Bjørlykke and Jahren, 2012). However, in deeply buried sandstones, grain-coating clay minerals have been shown to preserve porosity and permeability because they inhibit the growth of porosity-occluding quartz cement, leading to anomalously high reservoir quality (Ehrenberg, 1993; Taylor et al., 2010; Ajdukiewicz and Larese, 2012; Worden et al., 2020). Clay grain coats are micron-scale rims of clay minerals (e.g., chlorite) covering the surface of detrital grains, acting as a physical barrier to silica-rich fluids, preventing precipitation on a detrital quartz substrate (Xia et al., 2020). Clay grain coats can be detrital, formed in the sedimentary environment at the time of deposition, or authigenic, grown after deposition during diagenesis (Dowey et al., 2012); here we are interested in the origin and distribution of detrital grain coats that form in marginal marine environments.

Grain coating clays have been found to develop in all depositional environments but fluvially-influenced marginal marine environments such as deltas and estuaries have been identified as particularly efficient settings for chlorite-enriched sands (Dowey et al., 2012; Worden et al., 2020). This is suggested to be because of the influence of fluvial processes which increase deposition or generation of clay minerals, such as flocculation of clay aggregates due to the mixing of saline and fresh water, enhancing deposition of clays from suspension (Dowey et al., 2012; Worden et al., 2020). Core studies of ancient marginal marine reservoirs have found grain coat abundance to be facies-controlled (Ehrenberg, 1993; Martinius et al., 2005; Gould et al., 2010; Griffiths et al., 2021), and at the surface in modern analogue studies, sub-depositional environment has been found to be a control on detrital grain coat distribution (Wooldridge et al., 2018; Griffiths et al., 2019a; Wooldridge et al., 2019a), highlighting the importance in understanding sub-depositional environment/facies

organisation within these systems. The Ravenglass Estuary, northwest England, serves as a modern analogue for ancient and deeply buried marginal marine sandstone reservoirs due to the presence of detrital clay grain coats with a similar scale to some ancient and deeply buried reservoirs, and has been extensively studied. Therefore, it is an ideal testbed to assess the viability of machine learning, applied to geochemical and textural data, for sub-depositional environment classification, and its utility in understanding grain coat occurrence.

In this study, we will focus on the modern marginal marine sands of the Ravenglass Estuary to build on work by Muhammed et al. (2022) who first presented a basic machine learning approach to discriminate sub-depositional environment using sediment geochemistry, and Houghton et al. (2023) , who devised a more robust machine learning workflow using sediment textural attributes as predictors for sub-depositional environment. Here, we aim to expand the use of the machine learning workflow of Houghton et al. (2023) to incorporate geochemical data as predictors, including a new semi-automated approach to select the most informative elements for modelling, and ultimately assess the viability of this updated approach. Additionally, we will apply new texture and geochemistry-based models to classify sub-depositional environments in sediment cores from a modern tidal flat in the Ravenglass Estuary to offer an alternative interpretation of the evolution of the deposits when compared to manual core logging and point towards potential for reservoir quality-enhancing clay grain coat development by analogy to the estuary surface. Specific research questions addressed include:

1. What chemical elements are the most important when constructing geochemical classification models of the Ravenglass surface sub-depositional environments?

2. How accurate are machine learning models trained to classify sub-depositional environment using sediment geochemistry alone?

3. What effect does the integration of geochemical data with textural data have on the accuracy of machine learning models to classify sub-depositional environment?

4. What sub-depositional environments are preserved in the subsurface of a tidal flat in the Ravenglass Estuary and how are they distributed?

5. Can machine learning and geochemical data enhance prediction of reservoir quality in deeply buried sandstone reservoirs?

2. The Ravenglass Estuary

The Ravenglass Estuary is located in Cumbria, northwest England, UK, and has been studied to understand the distribution of clay minerals and generation of detrital clay grain coats as a modern analogue for deeply-buried marginal marine sandstones (Wooldridge et al., 2017a, 2017b; Griffiths et al., 2018; Wooldridge et al., 2018; Griffiths et al., 2019a, 2019b; Wooldridge et al., 2019b, 2019a; Simon et al., 2021; McGhee et al., 2022; Muhammed et al., 2022). The estuary is macrotidal, with a maximum tidal range of 7.55 m (Lloyd et al., 2013). In Ravenglass, the Rivers Irt, Mite, and Esk join to form a central basin that is protected by two barrier spits, ultimately flowing into the Irish Sea to the west (Fig. 1). The outer estuary, whose landward limit is defined by the transition from elongate channel-parallel sand bars to upper flow regime sand flats and tidal flats (Dalrymple and Zaitlin, 1994 ; Griffiths et al., 2018) (Fig. 1, Table 1), is dominated by wave processes but tidal processes dominate

environments upstream of the tidal inlet which define the inner estuary (Fig. 1C), with a strong flood tide flow and weaker ebb flow attributed to the shallow geometry of the estuary basin (Kelly et al., 1991).

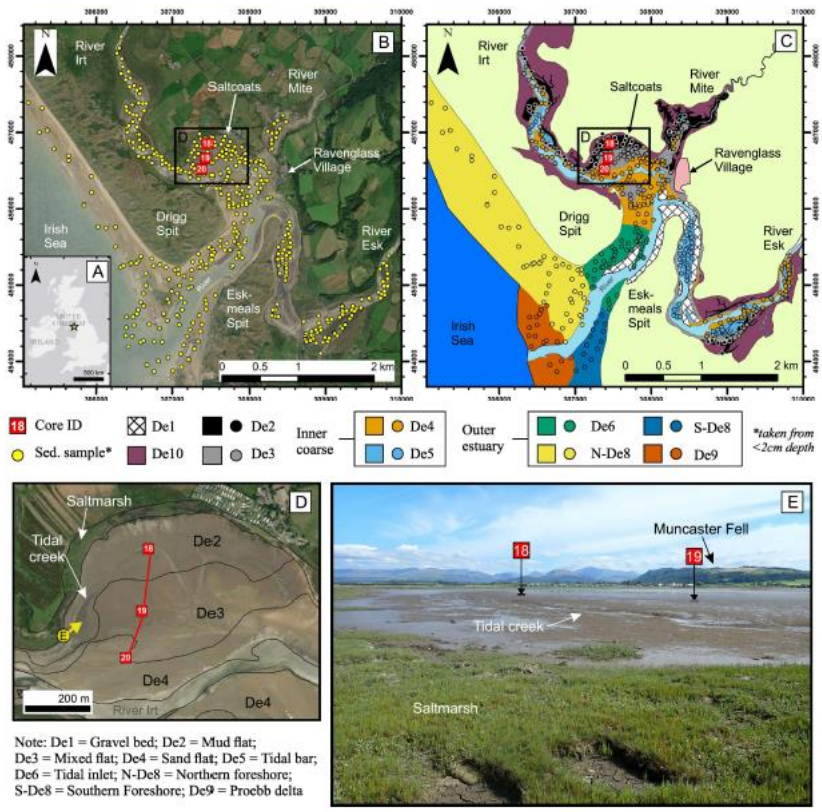

Fig. 1 A) Map of the United Kingdom with a gold star indicating the location of the Ravenglass Estuary; B) aerial imagery of the Ravenglass Estuary indicating locations of three geotechnical cores and 458 surface sediment samples (<2 cm depth); C) a map of present-day sub-depositional environments in the Ravenglass Estuary (modified after Simon et al., 2021; Houghton et al., 2023). All surface sediment samples are shown, coloured by their corresponding sub-depositional environment label. The river mapped here represents approximate low tide conditions; D) site map of Saltcoats tidal flat showing the location of geotechnical cores 18, 19, and 20 presented in this study (extent shown in B and C) and approximate location and direction of field photograph in E (yellow circle and arrow). Boundaries of sub-depositional environment are shown on top of aerial imagery; E) field photograph of the Saltcoats tidal flat during low tide showing approximate location of cores 18 and 19, with saltmarsh and tidal creek in the foreground.

Estuarine Zone (Model 2)	Sub-depositional environment (Model 1)	Description	Number of samples
N/A	Gravel bed (De1)	Deposits containing pebble-grade material (> 2cm), but may host all grain size fractions. containing andesitic, granitic, and sandstone lithologies. May contain disarticulated shell fragments.	0
Mud flat	Mud flat (De2)	15 – 50% sand (>63 μm), predominantly clay and silt settled from suspension, may be intensely bioturbated as indicated by surface worm casts and mottling.	55
Mixed flat	Mixed flat (De3)	Between 50-90% sand fraction (>63 μm) in the inner estuary, mixed deposition of suspended and bedload sediment.	94
Inner coarse	Sand flat (De4)	Inner estuary deposition of >90% sand fraction (>63 μm), may host ripples and/or low-amplitude dunes.	119
	Tidal bar (De5)	Prominent in-channel sand banks which are detached from the channel margin for all or some of the tidal cycle. May host dunes and ripples.	62
Outer	Tidal inlet (De6)	Area between barrier spits connecting the central estuarine basin to the open ocean, a transition from inner to outer estuary containing channel-parallel bars as a result of mixed wave and tidal processes.	43
	Backshore (De7)	Sediment lying above mean high water mark in the outer estuary, typically mix of wind-blown dune sand and marine sand.	-
	Northern foreshore (N-De8)	Stretch of sand-dominated deposits between low and high tide marks to the north of main channel, dominated by wave processes resulting in large wavelength/low amplitude dunes, upper phase plane beds and ripples.	51
	Southern foreshore (S-De8)	Similar to northern foreshore, but to the south of where the main channel meets the sea. Split from northern foreshore as they have significantly different texture and geochemistry (Simon et al., 2021; Muhammed et al., 2022).	17
	Proebb delta (De9)	Most distal portion of the estuary, a termination of the main channel where it meets the sea, depositing sand in a delta-like pattern.	17
N/A	Saltmarsh (De10)	Clay, silt and sand deposited at high tide with abundant rooted vegetation, typically flat top surface with eroded channel-facing bank.	0

Table 1. Summary of sub-depositional environments, including number of sediment samples and their corresponding 'estuarine zones'. Sub-depositional environments are used as classification labels in 'Model 1' and estuarine zones are used as classification labels in 'Model 2'. Gravel bed (De1) and saltmarsh (De10) can be visually discriminated and so no data are required for input into machine learning models.

Ten sub-depositional environments are identified in the estuary, each characterised for their texture, clay mineral content, clay coat coverage, and geochemistry at the surface (Wooldridge et al., 2017b; Griffiths et al., 2019b; Wooldridge et al., 2019a; Simon et al., 2021; Muhammed et al., 2022). The sub-depositional environments are as follows: gravel bed, tidal flats (here split into mud flats, mixed flats, and sand flats), tidal bars, tidal inlet, backshore, foreshore (split into northern and southern sub-environments), pro-ebb delta, and saltmarsh (see Table 1). Sub-depositional environments have also been grouped into 'estuarine zones' based on their texture and position within the estuary. Within these zones, mud flat (De2) and mixed flat (De3) remain separate, however sand flat (De4) and tidal bar (De5) have been merged to form the 'inner coarse' estuarine zone (Fig. 1C), and all outer estuary sub-depositional environments, tidal inlet (De6), both foreshore environments (N-De8 and S-De8) and pro-ebb delta (De9), have been merged to form the 'outer' estuarine zone (Table 1). This has been done as it is has been reported to improve classification accuracy when distinguishing the sand-dominated coarse inner estuary sediment from sand-dominated outer estuary (Simon et al., 2021; Muhammed et al., 2022; Houghton et al., 2023).

The sediments in the estuary are typically arkosic to subarkosic and host a typical mineral assemblage of quartz, plagioclase, K-feldspar, illite/muscovite, chlorite, kaolinite, and carbonate, however these minerals are not uniformly distributed as a function of sub-depositional environment or grain size (Griffiths et al., 2019a; Wooldridge et al., 2019a). Clay minerals (chlorite, illite, and kaolinite) and carbonates are most abundant in finer-grained inner estuary environments (mud and mixed flats), however it has been shown that clay minerals are also present within coarse lithic grains that are found in the outer estuary. Quartz abundance increases with grain size in the estuary and dominates sandy environments (Simon et al., 2021). These heterogeneities in mineralogy are also reflected in sediment geochemistry; for example, aluminium and potassium have been reported to be controlled by variations in illite abundance and calcium is strongly linked to the abundance of

calcium carbonate (calcite or aragonite) in shell fragments (Griffiths et al., 2019a; Muhammed et al., 2022).

The Saltcoats tidal flat is the largest tidal flat in the Ravenglass Estuary (0.35 km2), on the northern side of the River Irt channel as it enters the central basin (Fig. 1D, E). At the surface, all tidal flat sub-depositional environments are present: mud flat (<50 % sand fraction), mixed flat (50-–90 % sand fraction), and sand flat (>90 % sand fraction) (after Brockamp and Zuther, 2004; Wooldridge et al., 2018). Clay coat coverage in the Saltcoats tidal flat ranges between 2.5 % in sand flats and 76 % in mud flats, and has been shown to be strongly linked to sediment biofilms secreted by diatoms (Wooldridge et al., 2018). Underlain by Holocene peat and glacial diamicton, the estuarine sediment beneath Saltcoats is up to 3.4 m thick, and has been interpreted, by traditional facies analysis, to represent only mixed and mud flat deposits by McGhee et al. (2022).

3. Methods and materials

3.1. Surface sediment sample dataset

In this study, we have used the same surface sample set (where surface sediment is defined as being at < 2 cm depth) as Simon et al. (2021), Muhammed et al. (2022), and Houghton et al. (2023) supplemented with 11 additional samples from tidal bars, 18 additional samples from the tidal inlet, and 18 additional samples from the pro-ebb delta. In total, the dataset used here, to train machine learning models, consists of 458 surface sediment samples (Fig. 1, Table 1). Each sample was assigned a sub-depositional environment upon collection and transported to the University of Liverpool to obtain the particle size distribution using laser particle size analysis (LPSA), and the elemental composition using portable X-ray fluorescence (XRF). Attributes derived from the particle size distribution, obtained by LPSA, define the 'textural' component of the datasets, and the concentration of elements obtained by XRF analysis defines the 'geochemical' component of the datasets. These textural and geochemical data have been used to train a supervised machine learning model to classify sub-depositional environment in intervals of subsurface sediment from cores in the Ravenglass Estuary using equivalent data. The analytical and machine learning approaches are described in the following sections.

3.2. Geotechnical core collection and sampling

Three geotechnical cores were recovered from the Ravenglass Estuary in 2016, through the full thickness of the Holocene estuarine succession down to basal diamictons which have beenwere previously interpreted to be glacial till (McGhee et al., 2022). Core retrieval took place in one-metermetre segments using a Geotechnical 'P60' rotary rig, for use on harder sediment such as sand flats, or a Geotechnical 'Pioneer' percussion rig for softer sediment such as mud flats. To obtain textural data with LPSA, sediment samples (approximately 20 g) were extracted from the air-dried, cut face of the core at five-centimetre intervals, ensuring that only one lithological type was sampled for each interval. If a sampling interval fell on a sharp lithological boundary (e.g., a mud-sand contact) a sample was taken from each side of the boundary. Forty two sediment samples were taken from cores 18 and 19 each, and 22 taken from core 20, totalling 106 sediment samples for analysis by LPSA, coincident with intervals for analysis by XRF.

3.3. Bulk sediment texture and geochemical analysis

Sediment samples collected from the surface (n = 458) and subsurface (n = 106) of the Ravenglass Estuary were sieved at 2000 µm (to remove coarse material that is too large to be analysed by the LPSA instrument) and subjected to organic digestion with dilute hydrogen peroxide prior to grain size

analysis. Treated samples were analysed using a Beckman-Coulter LS13-320 laser particle size analyser and the raw data processed in GRADISTAT version 9.1 to obtain nineteen textural attributes (Blott and Pye, 2001). In this study we have selected the particle size distribution-derived textural attributes of mean grain size (µm), sorting (µm), primary modal grain size (µm), skewness, kurtosis, (after Folk and Ward, 1957) and volumetric proportions of very coarse sand, coarse sand, medium sand, fine sand, very fine sand, very coarse silt, coarse silt, medium silt, fine silt, very fine silt, clay, silt, sand, and mud as features for use in training textural machine learning models.

All sediment samples were analysed using an Olympus Vanta M-Series portable energy-dispersive XRF spectrometer. The instrument has a Rh anode X-ray tube, with analyses undertaken using a dual beam 'Geochem(2) + Cl' method involving a first beam scanning at 40 kV for 40 seconds and a second beam scan at 10 kV for 90 seconds. Scan times for each beam were selected by carrying out repeated scans on one part of the core, varying the scan time of each beam independently each time, and examining the reported errors associated with each beam scan time to identify when no significant improvement in error readings was achieved with increasing scan time (Supplementary data 1). The 'Geochem(2) + Cl' method is capable of reporting the concentrations and associated error of 36 elements between Mg and U, in parts per million (ppm), including 'Light Elements' (LE) to account for elements of lower atomic number than Mg that cannot be discriminated by the spectrometer, such as O, C, and Na. The elements measured by the XRF instrument for inclusion in training geochemical machine learning models are: LE, Mg, Al, Si, P, S, K, Ca, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, As, Se, Rb, Sr, Y, Zr, Nb, Mo, Ag, Cd, Sn, Sb, W, Hg, Pb, Bi, Th, and U. Loose surface sediment samples were air-dried for at least 24 hours and placed in a 25 mm diameter plastic dish with minimum 15 mm sample thickness for XRF analysis. Core sediment was analysed 'in-situ' on the dried cut surface of the geotechnical core in five-centimetre intervals, coincident with samples taken for textural analysis. In order to prevent contamination of the analysis window by the loose sediment but still allow full contact, surface samples and cores were scanned with an X-ray transparent 6 µm prolene film covering the sample/interval.

3.4. Adapted machine learning workflow

3.4.1. Workflow outline

This study employs the supervised decision tree-based Extreme Gradient Boosting (XGBoost) algorithm in RStudio (Chen and Guestrin, 2016; R Core Team, 2020) available through the 'xgboost' package and optimisation tools within the 'tidymodels' package (Kuhn and Wickham, 2020; Chen et al., 2022). The workflow employed to train models has been adapted from Houghton et al. (2023) who predicted sub-depositional environment (herein referred to as Model 1), and estuarine zone (herein referred to as Model 2) in the Ravenglass Estuary using sediment texture alone. The workflow has been adapted here to use transformed geochemical data, as well as textural data, as predictors for sub-depositional environment and estuarine zone. Outlined in Fig. 2, the machine learning workflow consists of four main sections (which correspond to code sections in Supplementary data 5, 6):

1. Importing and pre-processing of surface textural and geochemical data with sub-depositional environment labels.

2. Bayesian optimisation of XGBoost hyperparameters.

3. Training and testing classification models based on surface data with 4-fold partitioning.

4. Application of surface-calibrated classification models to classify sediment core intervals.

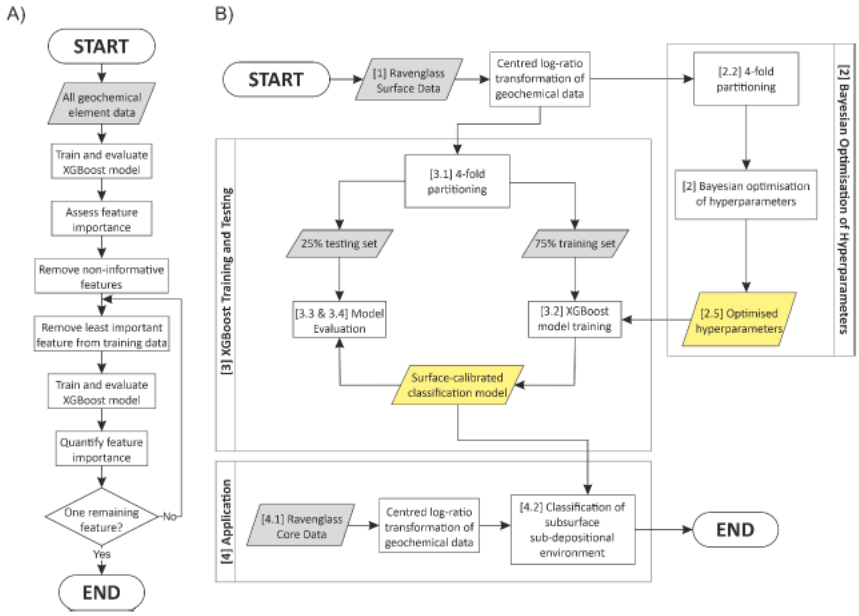

Fig. 2. A) Workflow of recursive feature elimination of geochemical elements to be completed before further model creation. At the end of this process the element combination with the highest mean overall accuracy is selected to be used in further geochemical models; B) overall machine learning workflow employed in this study to produce surface-calibrated classification models (adapted after Sun et al., 2020; Houghton et al., 2023). Bayesian optimisation (code section 2) must be completed before training and testing (code section 3). Numbers in square brackets refer to code section where the step occurs. Grey rhomboids indicate input data and yellow rhomboids indicate data generated internally by the workflow.

Hyperparameters control how a machine learning model is built, such as setting the maximum depth allowed in a decision tree. Default hyperparameter values are not suitable for all datasets or classification problems, and, if untuned, can lead to poor model performance, so it is necessary to find the optimum values suitable for each model. Here we use Bayesian optimisation (BO) and 4-fold cross-validation to select the optimum hyperparameter values (after Sun et al., 2020; Houghton et al., 2023). The BO process involves evaluating models made using different combinations of hyperparameters from within a search range set by the user. We have employed the search ranges of Houghton et al. (2023) but the upper limit of 'max_depth' and 'colsample_bytree' search ranges have been increased to account for additional input variables (i.e., the geochemical data) and increased model complexity (Table 2).

Model name	Search range (min, max)	No. input variables	Nrounds	Max_ depth	Min_child_ weight	Gamma	Eta	Colsample_bytree	Subsample	Overall Accuracy (%)
			10, 1000	4, 30	1, 20	0.1, 3	0.001, 0.1	2, no. input variables	0.6, 1	
Model 1-T		19	957	10	1	0.875903	0.003026	14	0.955686	67.05 (2.90)
Model 2-T		19	867	5	1	0.123283	0.011088	17	0.902921	83.43 (2.95)
Model 1-G		23	997	14	9	0.233562	0.091172	5	0.920026	57.90 (3.24)
Model 2-G		23	701	23	2	2.253253	0.082448	2	0.631333	82.98 (1.78)
Model 1-TG		42	654	17	4	0.213684	0.01754	22	0.619511	72.50 (2.22)
Mode 2-TG		42	519	20	6	0.572432	0.036395	6	0.800783	89.73 (1.91)

Table 2. Metrics used to evaluate models presented in this study. Each metric can vary between 0 and 1, where 0 is the 'worst' score, and 1 is the 'best' score. For 'Balanced Accuracy', a score of 1 indicates perfect agreement of classifications, a score of 0 indicates a perfect disagreement of classifications, and a score of 0.5 indicates a model performs as well as random chance. TN = 'True Negative', FN = 'False Negative', TP = 'True Positive', FP = 'False Positive'.

In order to evaluate the performance of individual classes within the machine learning models, we have employed the evaluation metrics: specificity, recall, precision, F1 score, and balanced accuracy, each designed to provide information about different aspects of model performance (Sokolova and Lapalme, 2009; Tharwat, 2020), these are summarised in Table 3. In addition, when discussing the overall performance of models, when considering all sub-depositional environments, we have used 'Overall Accuracy'.

Metric	Equation	Description
Specificity	$\frac{TN}{TN+FP}$	The effectiveness of a classifier at identifying true negative labels versus false positives
Precision	$\frac{TP}{TP+FP}$	The proportion of positive predictions that are true positives versus false positives
Recall	$\frac{TP}{TP+FN}$	The rate at which a classifier identifies true positive labels versus to false negatives
F1 score	$\frac{TP}{TP+\frac{1}{2}(FP+FN)} = 2.\frac{Precision x Recall}{Precision+Recall}$	A measure of accuracy combining precision and recall which does not take into account true negatives
Balanced accuracy	$\frac{Specificity+Recall}{2}$	The mean of specificity and recall, balancing the rate of true positives and true negatives
Overall accuracy	$\frac{\sum_{i=1}^{l}\frac{TP_i+TN_i}{TP_i+TN_i+FP_i+FN_i}}{l}$	Overall effectiveness of a classifier at identifying true labels, where 'l' is the number of classes

Table 3 Metrics used to evaluate models presented in this study. Each metric can vary between 0 and 1, where 0 is the 'worst' score, and 1 is the 'best' score. For 'Balanced Accuracy', a score of 1 indicates perfect agreement of classifications, a score of 0 indicates a perfect disagreement of classifications, and a score of 0.5 indicates a model performs as well as random chance. TN = 'True Negative', FN = 'False Negative', TP = 'True Positive', FP = 'False Positive'.

3.4.2. Compositional data analysis considerations

Geochemical data are compositional in nature and form closed datasets (i.e., they have a constant sum such as 1,000,000 ppm or 100 %), a property which may produce spurious results when carrying out statistical analyses on unprocessed data (Aitchison, 1982). In order to remove the effects of the closed dataset, data should be transformed using log-ratios of the components, which is particularly important when training machine learning models (Tolosana-Delgado et al., 2019). Thus, prior to using geochemical data to train machine learning models, raw elemental data (in ppm) were transformed using the centred log-ratio (clr) to remove the effects of a closed dataset (Eq. (1)) (Aitchison, 1982). This was achieved using the 'clr()' function in the Rstudio package 'compositions'; when using this function missing values are set to 0 (van den Boogaart and Tolosana-Delgado, 2021).

$$clr(x) = \left[\log \frac{x_1}{g(x)} \ldots \log \frac{x_D}{g(x)}\right] \quad (1)$$

Where: x = a sample of compositional data

xD = value of a single component in a sample

g(x) = geometric mean of all components in sample x.

3.4.3. Selecting geochemical input features: Recursive feature elimination

Although the XRF instrument is capable of reporting the abundance of 36 elements, some of these elements are not present above detection limits in the Ravenglass sediments for a large proportion of the samples analysed (e.g., Cd, Hg, U) (Supplementary data 2). If all of the elements detected by the instrument were included in training data, then 37 % of the dataset would contain missing values (i.e., below the limit of detection), that might adversely affect model performance by diluting more informative features. In order to determine the most informative geochemical elements to include in modelling, we have employed a 'recursive feature elimination' (RFE) process to allow us to eliminate elements that are either not used by the model or decrease model accuracy (Fig. 2A) (Chen and Jeong, 2007; Wang et al., 2020).

RFE was achieved by first training an XGBoost model on clr-transformed geochemical data of all 36 detectable elements to predict sub-depositional environments (i.e., Model 1) (code in Supplementary data 3). For this model the hyperparameter 'colsample_bytree' is set to a value 1, which makes all variables (geochemical elements) available for the model to use at all times, providing the maximum opportunity for the model to use any given element. All other hyperparameters are allowed to be optimised as normal. Then, the importance of each variable is quantified and ranked for the optimised model. This approach initially eliminates some variables (elements) which are never used by the model. Then, the least important element is removed, and a new model is created and evaluated using 4-fold partitioning, and the importance of variables within that model are assessed. This process is iterative, removing the least important elements one by one, re-assessing the least important variable between each successive model. Iterations continue until only one feature is left (Fig. 2A). The element combination which has the highest mean overall accuracy has been selected as the optimum features when training models with geochemical data.

4. Results

4.1. Geochemical feature selection

Assessing the importance of features in the first geochemical model that utilised all 36 elements showed that U, Sb, Sn, Cd, Ag, Se, and Cl are never used by the model. Therefore, these elements were not included in the following recursive feature elimination (RFE), or any further geochemical models, because their removal would have no effect on model performance. The results of the RFE on the remaining 29 geochemical elements are shown in Fig. 3, where the x-axis represents the model accuracy of training data from Si up to, and including, that element (from left to right), e.g., the 'Mn' position on the x-axis represents a model trained using all elements inclusive from Si to Mn. The optimum number of elements (highest mean overall accuracy) is 23, after the removal of Th, W, Hg, Mo, P, and Co, with an overall accuracy of 62.52 % (1sd = 8.12 %). The average accuracy of models remains above 60 % until there are only 9 elements remaining in the model (Si, Pb, Al, Ca, Mg, Zn, Sr, Mn, and S) suggesting that, even with limited geochemical data, reliable predictions can still be made. Removing elements Sr, Zn, and Mg has a small effect, remaining at approximately 55 % overall accuracy; however, beyond this, accuracy decreases rapidly with elemental removal until only Si is left as a predictor, which gives a mean overall accuracy of 38 %. The standard deviation decreases as more elements are removed from the model with values of approximately 8-–9 % when more than >20 elements are used by the model (see Y), decreasing to 4-–5 % below 16 elements (see Fe). Using elements up to and including Nb on Fig. 3 gave the highest mean overall accuracy, and so these elements have been selected to be taken forward into further geochemical models.

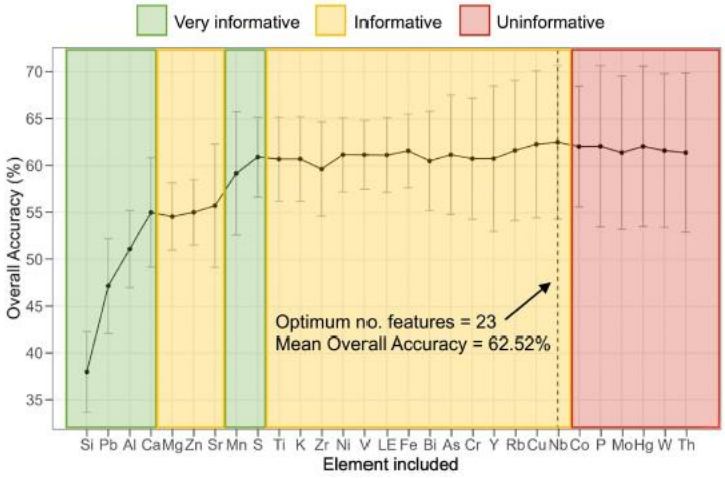

Fig. 3. Results of recursive feature elimination (RFE) for geochemical models showing the mean accuracy (black points) plus/minus one standard deviation of ten repeats of four folds (grey bars) versus each element added to the training dataset. Starting at Si, where a model is only trained on Si values with 38 % accuracy, moving to the right on the x-axis shows each element included in the training dataset. The highest mean accuracy is achieved with the inclusion of all elements between Si and Nb (inclusive) - these elements have been selected to be used in further geochemical models. U, Sb, Sn, Cd, Ag, Se, and Cl are not shown on this graph as they are never used by the initial model with all elements available, so their removal would have no impact on model performance. Regions of qualitative 'informativeness' are highlighted.

Chemical elements on Fig. 3 have been qualitatively labelled based on their informativeness, where 'uninformative' features (elements) have been removed from the modelling process as they decrease model accuracy, 'informative' features are elements included in training data but do not show significant improvement (<1.5 % increase) in model performance when they are included, and 'very informative' features show a significant improvement (>1.5 % increase) when they are included in training data.

4.2. Texture and geochemistry as predictors for sub-depositional environment in machine learning models

Six predictive models have been created using the outlined workflow (Fig. 4; Table 2; Section 3.4). 'Model 1' attempts to predict all eight discrete sub-depositional environments in the Ravenglass Estuary which cannot be visually discriminated (i.e., excluding gravel bed or saltmarsh as they can be visually discriminated; Table 1). 'Model 2' merges coarse-grained inner estuary environments (sand flat; De4, and tidal bars; De5), and outer estuary environments (tidal inlet; De6, northern foreshore; N-De8, southern foreshore; S-De8, and pro-ebb delta; De9) to produce four estuarine 'zones': mud flat (De2), mixed flat (De3), inner coarse (De4 + De5), and outer (De6 + N-De8 + S-De8 + De9) (Table 1). Models have been trained using 19 textural attributes obtained by LPSA and GRADISTAT processing (after Houghton et al., 2023) and clr-transformed geochemical data for 23 elements selected by recursive feature elimination (Fig. 3). The data used to train each model are indicated with a suffix, where '-T' is used to indicate models created using only textural data, '-G' is used to indicate models created using only geochemical data, and '-TG' indicates models created using combined texture and geochemistry. For example, 'Model 1-TG' refers to a model without merging of sub-depositional

environments trained on textural and geochemical data (Table 2). The results presented here use the same training/testing sets for model evaluation to eliminate potential sampling bias.

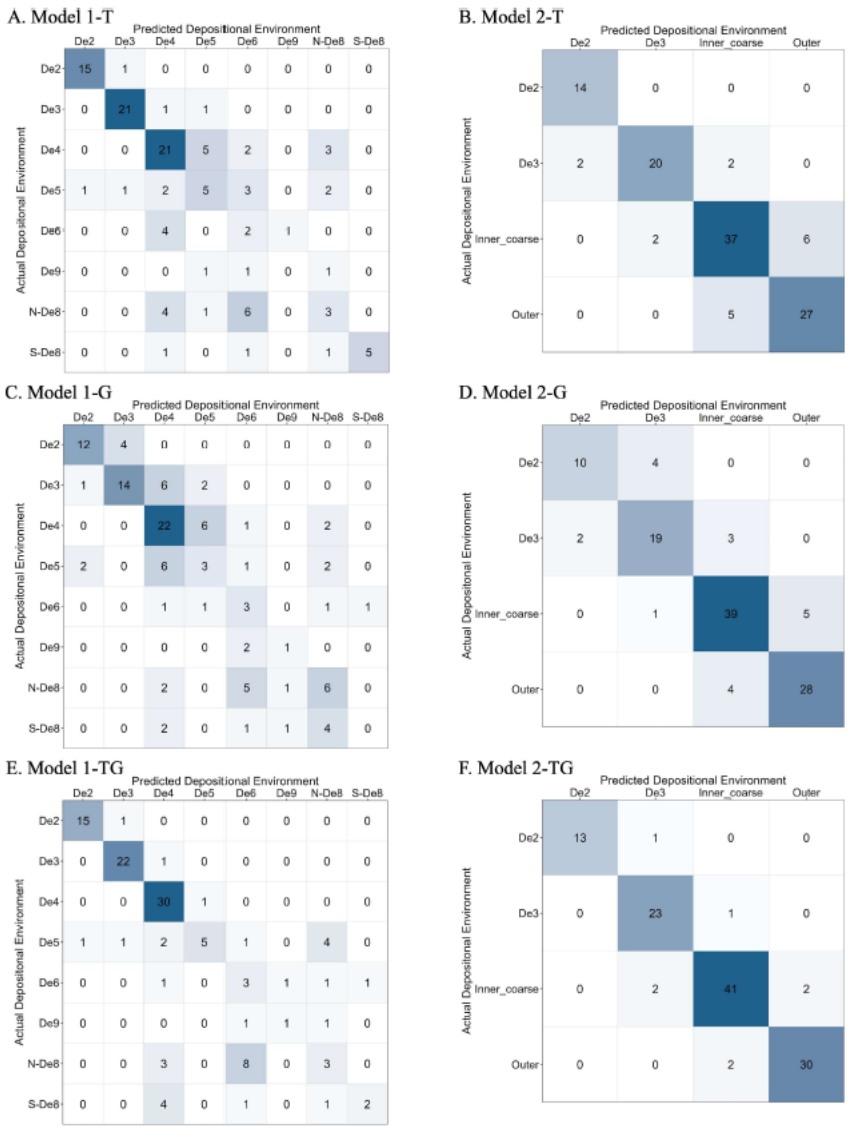

Fig. 4. Confusion matrices for six predictive XGBoost models for Model 1 (A–C, with no environment merging) and Model 2 (D–F, with environment merging). Model name suffix indicates training data, where: '-T' = textural, '-G' = geochemical, and '-TG' = combined textural and geochemical training data. Models in A and B were trained using 19 textural attributes (after Houghton et al., 2023), C and D show models made using clr-transformed geochemical data for 23 RFE-selected elements, and E and F contain models trained using both textural attributes and clr-transformed geochemical data. Matrices displayed for Models 1 and Model 2 represent the same respective training/testing sets. De2 = Mud flat; De3 = Mixed flat; De4 = Sand flat; De5 = Tidal bar; De6 = Tidal inlet; N-De8 = Northern foreshore; S-De8 = Southern Foreshore; De9 = Proebb delta.

4.2.1. Textural classification models (Model 1-T and Model 2-T)

The overall accuracy of Model 1-T is 67.05 % (1sd = 2.90 %), whilestand Model 2-T has an overall accuracy of 83.43 % (1sd = 2.95 %) (Fig. 4A, B). This is in contrast to the models made by Houghton et

al. (2023) who achieved an overall accuracy of 72.48 % and 85.32 % for Model 1 and 2 respectively. Model 1-T in this study shows improved recall scores for tidal bar (De5) and tidal inlet (De6) sub-environments compared to the equivalent model of Houghton et al. (2023) which indicates that, whilestalthough overall Model 1-T may have lower accuracy, it is more reliable for discriminating sandy sub-environments. Model 2-T in this study has a lower accuracy for the inner coarse zone as there are now two points of confusion with the mixed flat (De3) environment which was not observed by Houghton et al. (2023).

4.2.2. Geochemical classification models (Model 1-G and Model 2-G)

Geochemical models presented here were produced using only clr-transformed geochemical data of 23 RFE-selected elements (Figs. 3, 4C, D). Model 1-G has an overall accuracy of 58.07 % (1sd = 4.15 %) and Model 2-G has an overall accuracy of 82.98 % (1sd = 1.78 %). Compared to Model 1-T, mud flat (De2), mixed flat (De3), tidal bars (De5) and southern foreshore (S-De8) environments have lower recall scores whereas sand flat (De4), tidal inlet (De6) and northern foreshore (N-De8) have the same or greater recall scores (Fig. 4A, C). It is worth noting, however, that despite low recall scores of S-De8, it now has a very high specificity (0.991) meaning the model is highly unlikely to incorrectly classify samples from other environments as S-De8. Compared to Model 2-T, Model 2-G has lower recall scores for mud flat (De2) and mixed fat (De3) however there is small improvement in the inner coarse and outer estuary zones: from 0.822 to 0.867, and 0.844 to 0.875, reducing confusion with all other estuarine zones.

4.2.3. Integrated textural and geochemical classification models (Model 1-TG and Model 2-TG)

The overall accuracy for Model 1-TG is 72.50 % (1sd = 2.22 %) and for Model 2-TG is 89.73 % (1sd = 1.91 %). These accuracies show improvements over Models 1-T and 2-T by 5.45 % and 6.30 % respectively. Compared to Models 1-G and 2-G, overall accuracies are improved by 14.60 % and 6.75 % respectively. Comparing this model to separate textural and geochemical models shows improved, or the same, recall scores in all environments except S-De8 for Model 1 and improved or the same recall scores in all zones except for De2 in Model 2 (Fig. 4). Thus, a combination of textural and transformed geochemical data offers the best and most accurate option for classification of unknown samples. Models 1-TG and 2-TG have been taken forward for application to sediment cores as they offer the highest accuracy for their respective classification objectives; a summary of evaluation metrics for Model 1-TG is presented in Table 4 and for Model 2-TG in Table 5.

Sub-depositional environment	Specificity	Precision	Recall	F1	Balanced accuracy
Mud flat (De2)	0.992	0.955	0.939	0.944	0.966
($n_{test} \approx 16$)	(0.008)	(0.049)	(0.074)	(0.036)	(0.035)
Mixed flat (De3)	0.958	0.870	0.957	0.908	0.958
($n_{test} \approx 23$)	(0.034)	(0.082)	(0.029)	(0.039)	(0.010)
Sand flat (De4)	0.900	0.757	0.864	0.805	0.882
($n_{test} \approx 31$)	(0.029)	(0.041)	(0.064)	(0.028)	(0.026)
Tidal bar (De5)	0.964	0.705	0.468	0.550	0.716
($n_{test} \approx 13$)	(0.021)	(0.011)	(0.078)	(0.036)	(0.031)
Tidal inlet (De6)	0.943	0.466	0.419	0.423	0.681
($n_{test} \approx 7$)	(0.029)	(0.193)	(0.029)	(0.102)	(0.023)
Proebb deta (De9)	0.995	0.750	0.392	0.505	0.694
($n_{test} \approx 3$)	(0.005)	(0.250)	(0.180)	(0.205)	(0.092)
Northern foreshore	0.934	0.488	0.524	0.494	0.729
(N-De8) ($n_{test} \approx 14$)	(0.019)	(0.129)	(0.191)	(0.142)	(0.093)
Southern foreshore	0.984	0.538	0.629	0.508	0.807
(S-De8) ($n_{test} \approx 8$)	(0.007)	(0.175)	(0.266)	(0.107)	(0.131)

Table 4. Mean evaluation metrics of Model 1-TG for classification of sub-depositional environment using combined textural and geochemical data across 4-folds of testing data (1 standard deviation in parentheses).

Estuarine zone	Specificity	Precision	Recall	F1	Balanced accuracy
Mud flat (De2)	0.995	0.961	0.926	0.943	0.960
($n_{test} \approx 14$)	(0.005)	(0.039)	(0.054)	(0.044)	(0.029)
Mixed flat (De3)	0.964	0.875	0.947	0.909	0.955
($n_{test} \approx 24$)	(0.012)	(0.033)	(0.019)	(0.014)	(0.007)
Inner coarse	0.946	0.914	0.851	0.880	0.898
(De4 + De5)	(0.027)	(0.037)	(0.043)	(0.026)	(0.022)
($n_{test} \approx 45$)					
Outer	0.948	0.882	0.914	0.894	0.931
(De6 + De8 + De9)	(0.034)	(0.069)	(0.060)	(0.026)	(0.021)
($n_{test} \approx 32$)					

Table 5 Mean evaluation metrics of Model 2-TG for classification of estuarine zone using combined textural and geochemical data across 4-folds of testing data (1 standard deviation in parentheses).

4.3. Sub-depositional environment classification of core sediment

Three geotechnical cores taken from the Saltcoats tidal flat in the Ravenglass Estuary (Fig. 1) have been analysed for their texture and geochemistry to produce data equivalent to the surface samples that were used to train machine learning models (Supplementary data 4). This allows for the classification of each five-centimetre interval in core using the machine learning models. Here we have applied Model 1-TG and Model 2-TG, trained with textural and clr-transformed geochemical data (Fig. 4E, F, Tables 4, 5), to classify all core intervals; the predictions for cores are presented in Figs. 5, 6, 7 alongside a graphic lithology, and manual facies interpretation and radiocarbon dates (14C) from McGhee et al. (2022). Fig. 8 shows a correlation panel between cores 18, 19, and 20, influenced by the machine learning classifications to investigate the distribution of sub-depositional environments in the subsurface of the Saltcoats tidal flat.

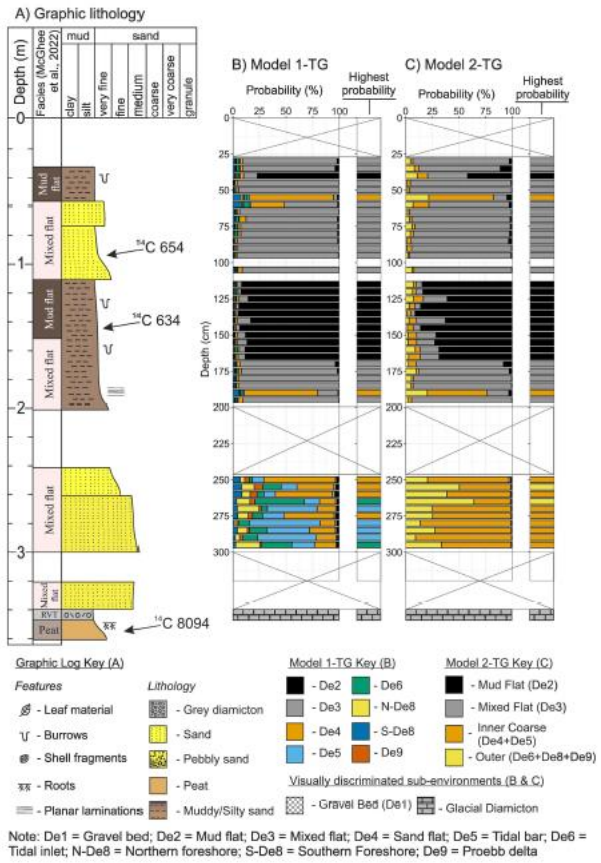

Fig. 5. Core 18 drilled through the mud flat (De2) of the Saltcoats tidal flat (see Fig. 1): A) graphic lithology log with facies interpretation of McGhee et al. (2022), and logs of classification probability and highest probability classification when using: B) Model 1-TG and C) Model 2-TG (14C dates from McGhee et al., 2022).

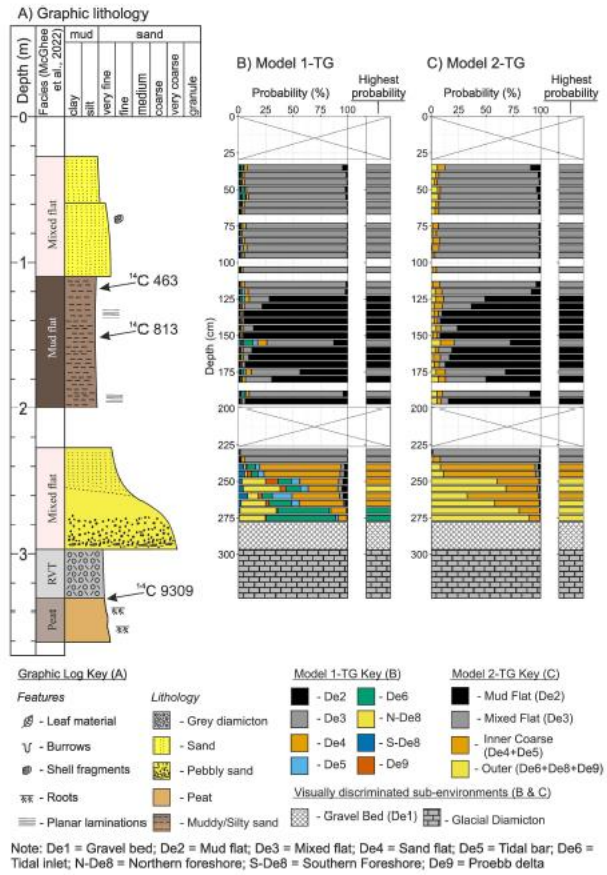

Fig. 6. Core 19 drilled through mixed flats (De3) of Saltcoats tidal flat (see Fig. 1): A) graphic lithology log with facies interpretation of McGhee et al. (2022), and logs of classification probability and highest probability classification when using: B) Model 1-TG and C) Model 2-TG (14C dates from McGhee et al., 2022).

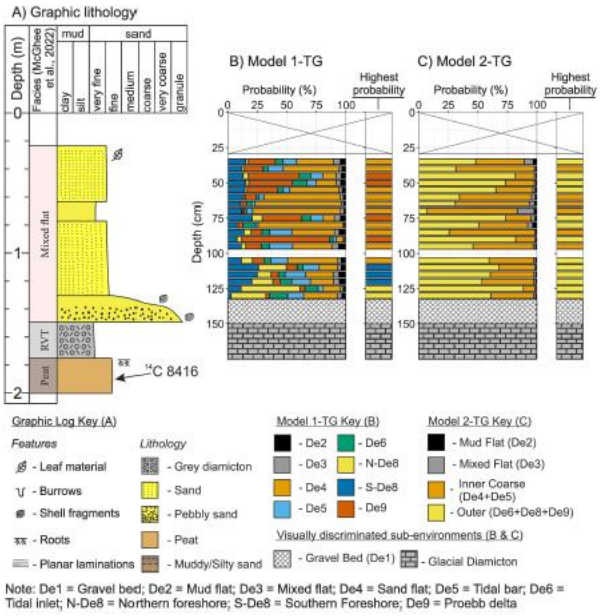

Fig. 7. Core 20 drilled through the sand flat (De4) of the Saltcoats tidal flat (Fig. 1): A) graphic lithology log with facies interpretation of McGhee et al. (2022), and logs of classification probability and highest probability classification when using: B) Model 1-TG and C) Model 2-TG (14C dates from McGhee et al., 2022).

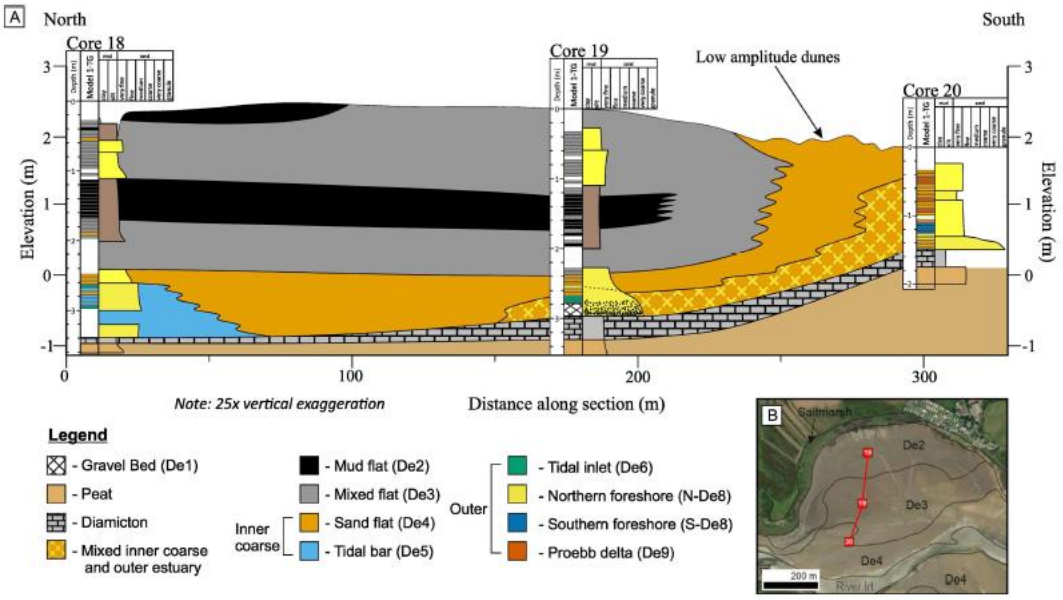

Fig. 8. A) Saltcoats tidal flat correlation panel between cores 18, 19, and 20 (see Figs. 5, 6, 7 for details). Shown on the panel are simplified graphic logs of each core and stick log of the most likely environment classified by Model 1-TG. Correlations have been informed by machine learning model output, and visual discrimination of peat and glacial diamicton which underly the estuarine deposits. The surface shown is the LiDAR topographic profile, allowing us to link subsurface correlation to surface mapped environments; B) map showing core locations and sub-depositional environment boundaries.

4.3.1. Core 18

Core 18 represents 3.4 m of estuarine deposition which is underlain by a thin layer (5 cm) of grey faintly laminated diamicton and peat (^{14}C 8,094 yrs BP) (McGhee et al., 2022). The first estuarine sediments in core 18, between 300 and 250 cm depth, are massive medium sand, classified by Model 1-TG as tidal bar (De5) and sand flat (De4), with two intervals of tidal inlet (De6), and by Model 2-TG as inner coarse with two intervals of outer estuary at 265 and 255 cm depth. Between 200 and 170 cm depth, sediment fines upwards with some planar laminations and is classified as mixed flat (De3) with one sand-rich interval highlighted as sand flat (De4). This De3 package fines upwards into mud flat (De2), that is 55 cm thick between 165 and 110 cm depth, which hosts some bioturbation-induced mottling. The top of this package is defined by a sharp lithological change from silt to very fine sand that is picked out in the graphic lithology at 110 cm depth. From 110 cm depth upwards, sediment is classified as mixed flat (De3). There is no consistent grading of grain size in this package but texture does vary internally, which defines 10 cm-scale beds represented in the graphic lithology.

4.3.2. Core 19

There is 2.95 m of estuarine sediment in core 19, underlain by grey diamicton containing root material and peat (dated at 9,309 ^{14}C yrs BP). The base of the estuarine succession (295-–275 cm) is classified as gravel bed (De1) as it predominantly contains material with a grain size greater than >2000 µm. Clasts in this section are up to pebble size (<45 mm) and sub-rounded, with predominantly andesite lithologies in a coarse sand and granule matrix. Above this, well sorted, medium sand fines upwards with 10 cm classified as tidal inlet (De6) by Model 1-TG and 15 cm of outer estuary by Model 2-TG. The fining upward trend continues above this package but its classification changes to inner estuary sand flat (De4). However, there is some disagreement with Model 2-TG which suggests that there is outer estuary sediment up to 250 cm depth, with only one interval of inner estuary coarse sediment. A lithological change occurs at 237 cm which cannot be picked out visually but is marked by transition from De4 to De3 classification. Above 237 cm, classification between Model 1-TG and 2-TG agree completely; sediment is predominantly classified as mixed flat (De3) with a mud flat (De2) package between 130 and 170 cm. Visually, sediment between 200 and 130 cm has a roughly uniform silty texture, however Model 1-TG and 2-TG pick out detail in differentiating intervals of mixed flat and mud flat. A sharp boundary at 115 cm depth marks the beginning of a fining upwards sequence containing two packages, distinguished by a surface separating a silt-rich upper layer from a silt-poor lower package. Despite visual differences in sediment composition here, all of these intervals from 120 to the top of the core are classified as mixed flat (De3).

4.3.3. Core 20

The estuarine sediment in core 20 is 1.5 m thick and is underlain by silty grey diamicton and peat (dated 8,416 ^{14}C yrs BP), similar to that seen in cores 18 and 19. The base of the estuarine sediment, between 150 and 130 cms, is classified as gravel bed (De1) and hosts disarticulated bivalve shells and angular clasts of andesite up to 30 mm in diametre. The top of the gravel bed is defined by a sharp lithological change at 127 cm depth, where sediment above is finer and more rounded. Fine sand deposits, between 130 and 105 cm depth, are classified by models 1-TG and 2-TG as foreshore (De8) and outer estuary respectively. Above 95 cm, however, classifications become less well-defined and inconsistent between models 1-TG and 2-TG. Model 1-TG classifies four of thirteen intervals as pro-ebb delta (De9) with all others as sand flat (De4), suggesting inner estuary deposition dominated. However, Model 2-TG classified seven of thirteen intervals as outer estuary and others as inner coarse, pointing towards a more balanced inner-outer deposition with a slight preference of outer estuary. There are no sedimentary structures present in this core to aid in the interpretation of these deposits.

4.3.4. Environmental correlation of the Saltcoats tidal flat subsurface

The three cores previously presented have been arranged into a scaled panel facing east to facilitate environmental correlations based on machine learning classifications; correlations are displayed with simplified graphic lithology and Model 1-TG classifications for each core (Fig. 8). Correlations between cores have been made based on Model 1-TG and 2-TG classifications connecting equivalent sub-depositional environments.

Figure 8 shows that diamicton and peat underly the whole Saltcoats tidal flat at approximately -–1 m elevation in cores 18 and 19, rising to 0 m elevation in core 20. Sand-dominated environments or a gravel bed, which typically fine upwards into the mud and mixed flats present at the surface sit, above the diamicton in each core. A thin correlatable package of mixed inner coarse and outer estuary sediment is interpreted to be present between cores 19 and core 20 with inner coarse above

this. In cores 18 and 19, the sequence of sand flat – mixed flat – mud flat – mixed flat is interpreted to be correlatable, forming horizontal bedding surfaces which terminate before reaching core 20, here illustrated by an interfingering morphology. Core 18 was recovered from the mud flat (De2) sub-depositional environment but this appears to be a thin veneer as it is not present in the top of the core, which begins at 30 cm depth. To remain consistent between surface and subsurface environments, a thin package of mud flat is indicated on the tidal flat surface which terminates about half-way between cores 18 and 19. The tidal bar (De5) package present at the base of core 18 is not considered to be correlatable with other cores, and is shown to interfinger with the overlying and adjacent inner estuary sand flat (De4) because, at the surface, tidal bars are seen to be exclusively adjacent to inner estuary environments, rather than interfingering with the outer estuary intervals at the base of core 19. It is noteworthy that this machine learning-led reinterpretation of the cores, based on a combined geochemical- and textural-informed model, has revealed a rich image of the depositional evolution of this apparently simple area over the last ~ 9,000 years.

5. Discussion

5.1. Geological links between geochemical elements and feature importance

Recursive feature elimination revealed that of 29 elements used by the model, just five elements account for 22.2 % of the 24.5 % overall increase in accuracy from Si (the most important element); these are highlighted as 'very informative' (green) on Fig. 3; these five elements are Pb, Al, Ca, Mn, and S. Si, Al, and Ca are major mineral-forming elements, with Si and Al present in the Ravenglass estuary as quartz, feldspars, and clay minerals (i.e., chlorite, kaolinite, and illite) (Griffiths et al., 2019a). Considering this, it is possible that Si and Al may be used to distinguish Si-rich/Al-poor sediments of coarse -grained environments from Si-poor/Al-rich fine -grained environments which host a greater proportion of Al-bearing clay minerals (Van Hoang et al., 2010; Muhammed et al., 2022). Lead, perhaps surprisingly, is shown to be the second most important element used by the XGBoost model to distinguish sub-depositional environments, leading to a 9.18 % increase in overall accuracy when included with Si (Fig. 3). Lead mineralisation (as galena, PbS) has been widely reported across Cumbria, present in mineral veins also associated with pyrite and other sulphide minerals containing Cu, Zn, and As, elements which are also shown to be 'informative' after recursive feature elimination (Cameron et al., 1993) (Fig. 3). Calcium (3.91 % increase from Al) is mainly present in the estuary in carbonate minerals as detrital grains (i.e., dolomite derived from the nearby Lower Triassic Sherwood Sandstone) and shell material (Griffiths et al., 2019a). Muhammed et al. (2022) reported that Ca concentration was highest in mud flat, mixed flat, and southern foreshore environments, reflecting the distribution of carbonate minerals reported by Griffiths et al. (2019a). The simple RPART classification model of Muhammed et al. (2022) used the index K/(K + Ca) to partition samples of northern foreshore from sand flat and tidal bar, and the index Ca/(Ca + Fe) to partition samples of southern foreshore from pro-ebb delta. Mn (3.46 % increase from Sr) is not a major constituent of any common minerals in the estuary, but instead may be present as a trace element in detrital lithic grains containing chlorite or biotite, possibly substituting for Fe due to similarities in geochemical properties. Sulphur is also shown to have a significant impact on model performance (1.74 % increase from Mn), and is interpreted to be present in the sediment mainly as iron sulphide, with pyrite (FeS_2) as the main identified sulphide mineral present in the estuary (Griffiths et al., 2018). Pyrite in the estuary surface is restricted to mud and mixed flat environments as framboids which may be embedded in clay grain coats, interpreted to form by bacterial sulphate reduction of marine aqueous sulphate (Berner, 1980; Griffiths et al., 2018).

In the broadest sense, the bulk sediment input into an estuary reflects its hinterland geology and local weathering processes, ultimately affecting the mineralogy and geochemistry of sediment within the estuary (Garzanti et al., 2009; Caracciolo, 2020). The results of geochemical modelling therefore reflect not just hydrodynamic processes, and the resulting organisation of sub-depositional environments (which are intrinsic processes in marginal marine systems), but also aspects of the specific system that are extrinsic, i.e.e.g., sediment provenance. The prominence of lead (Pb) in geochemical modelling here highlights the issue of hinterland-specific provenance as lead occurrence is likely to be due to the presence of lead mineralisation in the hinterland (Cameron et al., 1993). Without lead mineralisation in the hinterland, it is possible that lead levels in the sediment would be below detection and therefore not useful for the classification models. The interaction of physical grains and estuarine hydrodynamic processes results in the observed heterogeneity in sediment texture; consequently, the observed heterogeneity in geochemistry is largely a function of how extrinsic factors affect how mineralogy changes with grain size rather than directly influencing hydrodynamic interaction (Boyd et al., 2006; Garzanti et al., 2009; Ainsworth et al., 2011). Therefore, it is suggested that applying geochemical classification models developed for one hydrodynamic basin to another is probably inappropriate and likely to be less accurate than applying textural models as sediment texture reflects processes which are intrinsic, whereas geochemistry is strongly controlled by provenance, possibly introducing informative system-specific elements into models, as is seen in the Ravenglass Estuary.

5.2. Classification model performance

In this study we have employed the general workflow of Houghton et al. (2023) and adapted it to semi-automatically select informative geochemical elements for inclusion in training data in the same sedimentary system, the Ravenglass Estuary, to predict sub-depositional environment using combined sediment texture and geochemistry. Simple recursive partition classification schemes attempts were also developed by Simon et al. (2021) and Muhammed et al. (2022). Therefore, we can make direct comparisons to these models, and discuss where and why the performance of textural and geochemical models may differ.

Model 1-T presented here, compared to the results of Houghton et al. (2023), has improved recall scores for Tidal bars (De5; 0.468 vs 0.214) and Tidal inlet (De6; 0.524 vs 0.0). The difference between these two models is the addition of new datapoints sampled from De5 and De6, and slight expansion of hyperparameter search ranges. With additional training/testing data, these two previously poorly predicted classes may have passed a threshold to become defined enough within the machine learning model so that their accuracy has increased. However, one effect of this additional sampling has also been to decrease the specificity scores of both De5 and De6 environments, meaning that these classes are now more susceptible to false positives (Fig. 4A).

Model 1-G and Model 2-G have mean overall accuracies of 57.90 % and 82.98 % respectively. Muhammed et al. (2022) were able to achieve overall accuracies 72.3 % and 87.0 % for equivalent models using the decision tree-based (but not gradient boosted) Recursive Partitioning and Regression Trees (RPART) algorithm trained using indices of selected geochemical elements. The decreased accuracies presented here for Model 1-G and 2-G in comparison may be because XGBoost has reduced overfitting, and is sensitive to outliers, however the RPART model may be overfitted, producing a model that is more accurate but less representative. Additionally, in this study we have supplemented the original dataset with samples from sub-depositional environments which have been shown to perform poorly in classification models (e.g., pro-ebb delta, tidal bars, and tidal inlet) in order to achieve a more balanced dataset, resulting in the models appearing to perform worse overall purely as a result of additional sampling (Simon et al., 2021; Houghton et al., 2023).

Muhammed et al. (2022) selected their geochemical element indices used for model training by analysis of variance (ANOVA) and honestly significant difference (HSD) tests, which provided the justification for splitting of sub-depositional environments based on geochemistry. We have built on this approach by allowing potentially all of the geochemical data to be used, rather than a limited number of indices which may not capture all informative features and demonstrated a semi-automated approach for element selection by recursive feature elimination that improves model performance by removing uninformative features.

Combining textural and geochemical data to train models improves classification accuracy when compared with independent datasets (Fig. 4). With both data types available in the training process, the most informative features can be exploited further in more dimensions. This means it is possible that the model can use the pre-defined textural differences between tidal flat sub-depositional environments (Table 1) whilest simultaneously being able to recognise the sand-dominated sub-depositional environments, such as De4, De6, and N-De8, which are better distinguished using geochemistry in Model 1-TG (Fig. 4A, C & E) and reducing confusion between inner coarse and outer zones in Model 2-TG (Fig. 4B, D, F). However, this effect is not seen in all sub-depositional environments; in Model 1-T, five of eight S-De8 samples are correctly classified, yet this is not carried forward into Model 2-TG where only two samples are correctly classified (Fig. 4F). This may be because, despite an apparent strong distinction when using texture, there is a contradicting geochemical signal which introduces uncertainty which that cannot be resolved by the modelling process.

5.3. The history of deposition in the Saltcoats tidal flat

Estuarine sediment in the Saltcoats tidal flat is underlain by diamicton, which is in turn underlain by peat. The grey diamicton, described at the base of the estuarine sediments were was interpreted by McGhee et al. (2022) to be part of the Ravenglass Till Member (RVT). The RVT was described by Merritt and Auton (2000) as a massive brown subglacial diamicton with metre-scale thickness deposited during the Late Devensian, approximately 20,000 yrs BP. The diamicton observed beneath Saltcoats, at approximately 3 m depth, is grey and locally contains laminations and/or organic material. The peat that immediately underlies the diamicton was dated reported by McGhee et al. (2022) to sit between 8,094 and 9,309 14C yrs BP in core thereby constraining the age of the diamicton. These features and dates are not consistent with the description or timing of the RVT. It is instead suggested that the thin diamicton deposit here is more likely to represent post-glacial lacustrine clay-peat deposits, similar to those reported by Smith et al. (2022) and Coleman et al. (2021) in other parts of west Cumbria.

The machine learning classification output (from Model 1-TG and Model 2-TG) provides an unbiased interpretation of the subsurface estuarine deposits in the Saltcoats tidal flat as an alternative to qualitative core description presented by McGhee et al. (2022). Across the three cores, all eight sub-depositional environments are classified by Model 1-TG (Figs. 5, 6, 7). However, intervals classified as pro-ebb delta or southern foreshore (as in core 20; Fig. 7B) are unlikely to actually represent deposition in these specific sub-environments due to confusion with other sand-dominated environments based on surface training/testing data (Fig. 4E). However, tThe classification of outer estuary sediment in these same intervals by Model 2-TG (Fig. 7C) is a more reliable classification than specific sub-depositional environment classifications made by Model 1-TG, as indicated by the minimal confusion with inner coarse samples and very high evaluation metrics (Fig. 4E; Table 5).

Significantly, outer estuary intervals are classified at the base of the estuarine sediments in the present inner estuary (e.g., cores 19 and 20; Figs. 5, 6, 7). This suggests that the shoreline was set

back (further east) by up to 1.4 km relative to its current position and the site has since transitioned to the inner estuary. This interpretation is consistent with the Holocene transgression and marine incursion which occurred prior to the formation of the central basin. Subsequent coastal progradation during gradual sea level fall sheltered the area from the high energy, wave-dominated environment of the outer estuary by the formation of the Drigg Spit (Lloyd et al., 2013; McGhee et al., 2022). Upon transition to a relatively quiescent, inner estuary environment, conditions appear to have remained similar to those found at the present time such as to allow the deposition of mud and mixed flat seen across the present day Saltcoats surface. This points towards tidal processes and deposition from suspension dominating across the Saltcoats tidal flat since the transition to inner estuary.

The basal estuarine sediments in core 18 (Figs. 5, 8) have been classified as tidal bar (De5), which given the high specificity score of 0.964 (Table 4), suggests that this is a reliable classification. This interpretation is supported by an inner coarse classification by Model 2-TG (Fig. 5). Presently in the estuary, tidal bars form in a channel or immediately adjacent to a channel, suggesting that a river channel once ran under the tidal flat at this location. The orientation of this channel cannot be determined based on the available data; however, it may represent an early northeast-flowing diversion of the Irt following the arcuate shape of the tidal flat saltmarsh margin, or a westward flowing diversion of the Mite. Fig. 9 shows a possible spatial evolution of sub-depositional environments over time in the Saltcoats tidal flat, based on the environmental interpretation of the correlation panel by machine learning classification and knowledge of the previous highstand (~6,500 yrs BP) and subsequent gradual sea level fall within the estuary (Fig. 8) (McGhee et al., 2022).

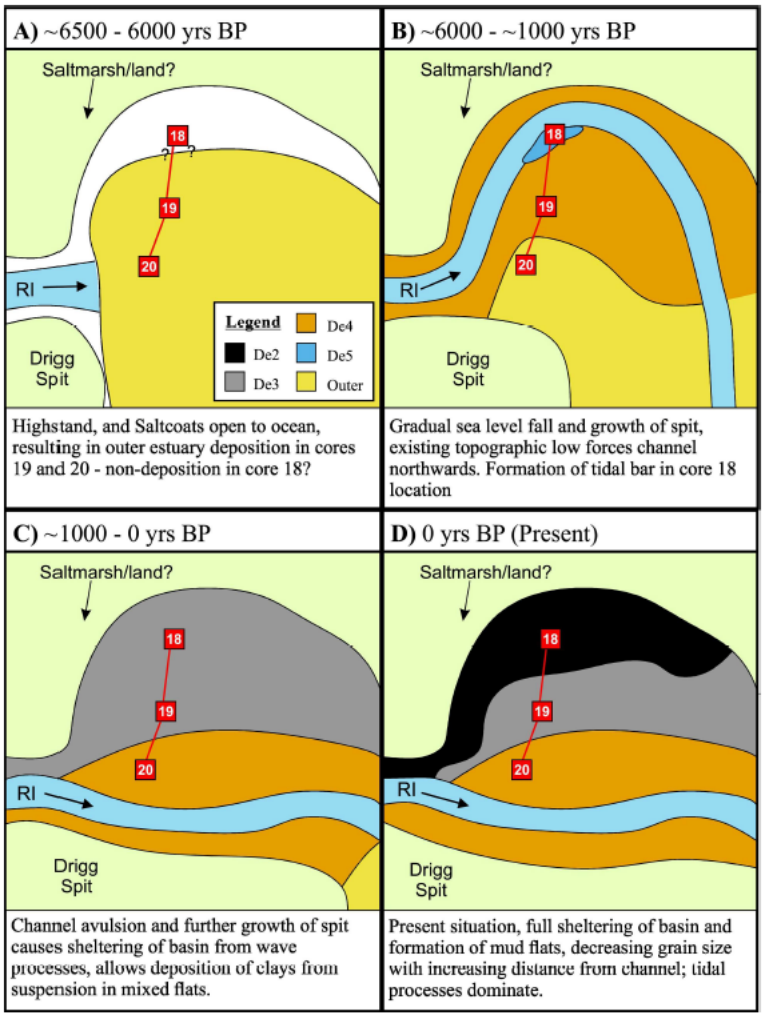

Fig. 9. Schematic diagram of possible local evolution of estuarine sedimentation in the Saltcoats tidal flat. Four time steps are shown: A) post-glacial Holocene highstand (<6000 yrs BP; McGhee et al., 2022) prior to spit development; B) partial spit development and sea level fall cause outer estuary to retreat; C) advanced spit development and sheltering of Saltcoats from high energy wave action; D) present-day environment (RI = River Irt input; De2 = Mud flat; De3 = Mixed flat; De4 = Sand flat; De5 = Tidal bar).

5.4. Implications for Ppredicting estuarine sandstone reservoir quality

Conventionally, high reservoir quality can be found in well-sorted, quartz-rich, clay-poor sediments (Beard and Weyl, 1973). In the Ravenglass Estuary, coarse inner estuary (sand flats and tidal bars) and outer estuary environments are well sorted with the least clay-grade material (Wooldridge et al., 2018; Griffiths et al., 2019a). Fine-grained inner estuary environments (mud and mixed flats) are poorly sorted and contain more than >10 % mud-grade material by volume which is likely to lead to conventionally poor reservoir quality, with mud flats experiencing enhanced ductile mechanical compaction due to a greater abundance of ductile clay minerals (Worden and Utley, 2022).

With the onset of quartz cementation upon deeper burial (>2 km depth at > 70 °C), quartz-rich sediments (such as coarse inner and outer estuary environments) may develop extensive quartz cement, destroying porosity and permeability (Worden and Burley, 2003; Ajdukiewicz and Lander,

2010). In finer-grained, inner estuary deposits, however, reservoir quality may be preserved due to the development of diagenetic clay grain coats, acting to inhibit the development of quartz cementation, and formed by the transformation of detrital clay coat precursors (Griffiths et al., 2019a ; Virolle et al., 2019a; Griffiths et al., 2021). It has been reported that a clay mineral volume between 3.5 and 13.0 % is sufficient to form continuous grain coats, termed the 'Goldilocks zone' by Wooldridge et al. (2017b) because, in this range, there is enough clay material to effectively inhibit quartz cementation but not so much as to block pore throats, which would negatively affect reservoir quality. In the Ravenglass Estuary, the Goldilocks zone comprises central basin tidal flats, and upper estuary tidal bars (Wooldridge et al., 2017b), where the equivalent environments in this study are mud flat (De2), mixed flat (De3), and tidal bars (De5). Wooldridge et al. (2019a) reported mud flat clay content varied from 8.7 to 17.8 % indicating, in some cases, that mud flats contained too much clay to preserve reservoir quality. Conversely the clay content of tidal bars (vegetated and non-vegetated) varied from 2.4 to 8.0 % indicating that some samples did not contain enough sufficient clay to lie within the Goldilocks zone (Wooldridge et al., 2019a). In the case of mud flats, however, clay content varied from 3.8 to 13.7 %, indicating that most samples sat within the Goldilocks zone,; it is therefore suggested that mixed flats (De3) are the optimum target for reservoir quality-preserving clay grain coats and will be the focus for further discussion. Using our combined approach of textural and geochemical analysis with machine learning, we have shown that the environments within the Goldilocks zone can be reliably classified by the models in the case of mud flats and mixed flats, but is less reliable for tidal bars (Fig. 4E, F).

The traditional interpretations made by McGhee et al. (2022), suggested that these cores are dominated by mixed flats with minor mud flats, an interpretation which might suggest it hosts grain coats and thus has a high potential for enhanced reservoir quality upon deep burial. However, the automated machine learning classification model presented here suggests that the deposits are much more varied, containing coarse inner estuary sand flats and outer estuary sediment as well as mud and mixed flats. For example, the basal package of each core presented here was interpreted to be 'mixed flats' by McGhee et al. (2022), however, Model 2-TG labels them as coarse inner (in core 18; Fig. 5) or outer estuary sediment (in cores 19 and 20; Figs. 6, 7) using Model 2-TG, suggesting limited potential of grain coat formation and high susceptibility to quartz cementation upon deep burial (Griffiths et al., 2019a). Overall, the classification of Model 1-TG agrees with the interpretation of McGhee et al. (2022) in 24/42 intervals in core 18 (Fig. 5), 25/42 intervals in core 19 (Fig. 6), and 0/22 intervals in core 20 (Fig. 7). Across all three cores, most disagreements occur in sand-rich intervals, representing 41 of 57 disagreements, with remaining 16 out of 57 disagreements in the definition of mud flat (De2) and mixed flat (De3) intervals. This outcome highlights how the classifications made by Model 1-TG and Model 2-TG are able to discriminate packages of mixed flat from inner estuary sand flat and outer estuary environments in the core with a direct link to surface properties, where a standard facies approach would not be capable of this subtle differentiation. The ability to distinguish sub-depositional environments in such a way in core has further implications on understanding the sub-depositional environment-related variability of reservoir quality upon burial and diagenesis (Griffiths et al., 2019a). Here we have found that mixed flats are abundant in the subsurface of the Saltcoats tidal flat (Figs. 5, 6, 8), and constitute 43 % (18/42) of classified intervals in core 18 (Fig. 5), and 48 % (20/42) of classified intervals of core 19 (Fig. 6). These proportions suggest that sheltered tidal flats are a stable environment for the long-term accumulation and preservation of Goldilocks zone mixed flat sediment as well as forming a really extensive deposits which may host enhanced porosity and permeability upon deep burial due to the inhibition of quartz cementation by clay grain coats (Fig. 1C). However, thicker mixed flat deposits may also be associated with mud flat baffles, as seen in cores 18 and 19, potentially limiting connectivity of these deposits (

Fig. 8). The subsurface classifications using machine learning models can provide basic information on grain coat potential by analogy to equivalent surface sub-depositional environments however these links require confirmation by mineralogical and grain coat coverage analysis; here we provide a framework to investigate these aspects in the subsurface.

Results presented here show that the inner estuary mixed flats (De3) have high evaluation metrics, with a minimum of 0.786 across all metrics (Table 6). Values to highlight are specificity (Table 3), which, across all models, has a minimum value of 0.948, indicating that samples which are not mixed flat are rarely incorrectly labelled as mixed flat. For mixed flats, Models 1-T and 2-T perform highly across all metrics, and this is expected due to the original texture-based definition of mixed flats (between 50 and 90 % sand fraction) which are identified easily by the models and exploited. For Models 1-G and 2-G, however, there are no pre-defined geochemical characteristics that can be used to discriminate mixed flat from other environments. As such, performance metrics of geochemical models are lower than the textural models but still remain high, with recall values (the rate of true positive prediction versus false negatives) of 0.786 and 0.799, and mean F1 scores (a measure of accuracy combining precision and recall which does not take into account true negatives) of 0.786 to 0.810, indicating that De3 samples are reliably identified, and that their accuracy is not inflated by an imbalance of true negatives (Tables 3 & 6). Importantly, specificity values remain very high (0.948 and 0.956 for Models 1-G and 2-G respectively) and comparable to those of the textural models (Table 6). These results suggest that geochemistry is a reliable predictor for the mixed flat environment, and therefore potentially suitable for identifying sediments with a high likelihood of containing reservoir quality-enhancing grain coats.

Model name	Specificity	Precision	Recall	F1-score	Balanced accuracy
Model 1-T	0.995 (0.005)	0.969 (0.031)	0.939 (0.074)	0.951 (0.037)	0.967 (0.036)
Model 2-T	0.961 (0.016)	0.867 (0.046)	0.937 (0.062)	0.898 (0.026)	0.949 (0.026)
Model 1-G	0.948 (0.022)	0.800 (0.048)	0.786 (0.119)	0.787 (0.063)	0.867 (0.054)
Model 2-G	0.956 (0.017)	0.823 (0.071)	0.799 (0.103)	0.810 (0.086)	0.878 (0.059)
Model 1-TG	0.958 (0.034)	0.870 (0.082)	0.957 (0.029)	0.908 (0.039)	0.958 (0.010)
Model 2-TG	0.964 (0.012)	0.875 (0.033)	0.947 (0.019)	0.909 (0.014)	0.955 (0.007)

Table 6 A comparison of mean evaluation metrics for the Mixed Flat (De3) sub-depositional environment, most likely to host the optimum amount of clay minerals for grain coat generation, across 4-folds of testing data (1 standard deviation in parentheses) for all six XGBoost classification models presented (Fig. 4).

5.5. Machine learning as a tool for aiding reservoir quality prediction

Machine learning is being increasingly explored as a tool to aid reservoir quality prediction in oil and gas fields, and can be utilised in any geological setting, with almost any data type (Saporetti et al., 2019; Yasin et al., 2021; Hansen et al., 2023; Liu et al., 2023). Wireline log data have been used as predictors for diagenetic facies or litho-facies due to the ubiquitous acquisition of wireline data and its utility in distinguishing sediment attributes and reservoir properties even without machine learning techniques (Dubois et al., 2007; Xie et al., 2018; Sun et al., 2020; Deng et al., 2021). In the tidal flat-dominated Longtan Formation of the Sichuan Basin, Liu et al. (2023) reported that wireline

data can be used to predict eight lithofacies classes with recall values ranging between 0.60 and 1.00 when using a Random Forest model, a supervised decision tree-based algorithm comparable to XGBoost. The performance of Model 1-TG in this the present study has comparable recall scores for mud flats, mixed flats, sand flats, and both foreshore environments (Table 4) whereas tidal bar, tidal inlet, and pro-ebb delta have lower recall scores due to confusion with other sand-dominated environments because of their textural and geochemical similarity (Fig. 4A, C, E).

Hansen et al. (2023) employed a supervised machine learning approach using wireline log data as predictors for core plug helium porosity for regional prediction of reservoir quality in the Stø Formation, SW Barents Sea. Based on petrographic evidence, illite grain coats were concluded to have inhibited the growth of quartz cement and are considered to be the most important control on reservoir quality in the formation (Hansen et al., 2017; Løvstad et al., 2022). Previous studies of the illite grain coats focused on thin section analysis and core plug helium porosity, however Hansen et al. (2023) highlighted the need for core, and limited resolution and discontinuous nature of the data as downsides of these techniques.

The approach outlined in this study has demonstrated that sub-depositional environments within the 'Goldilocks zone' of grain coat coverage can be reliably predicted using geochemistry and textural data. Further work could attempt to translate sub-depositional environment-based near-surface studies and machine learning approaches into subsurface deposits to predict facies that host reservoir quality-enhancing grain coats. The workflow used in this study could be used as an additional tool when conducting reservoir studies, for application to XRF data obtained from core, or down-hole geochemical logging tools (Herron and Herron, 1990). Post-depositional diagenetic processes may act to alter geochemical or textural attributes in lithified sediment which could affect predictions of depositional facies and so these processes should be understood before attempting to employ the approach we present (Morad et al., 2000). In this study we have chosen sub-depositional environments as the classification labels, however this could be adapted to predict facies (if reservoir quality is facies-controlled), or it may be adapted to a quantitative regression problem to predict clay coat coverage or porosity. Such models would allow for high resolution and continuous interpolation of reservoir core, providing information to improve reservoir models. This could include highlighting intervals of high or low reservoir quality, or providing information on the distribution of grain coats in a core (if they are present).

6. Conclusions

In this study, textural and geochemical data from the surface of a modern macro-tidal estuary (the Ravenglass Estuary, northwest England) have been used as predictors for sub-depositional environment and estuarine zone using an established machine learning (ML) workflow utilising the Extreme Gradient Boosting algorithm in RStudio. Six surface-calibrated classification models have been trained and evaluated to explore the effectiveness of different data types in creating machine learning classification models. The most accurate models (integrated geochemistry and texture) were used to classify sub-depositional environment for equivalent data obtained from five-centimetre intervals of three cores drilled through a tidal flat in the Ravenglass Estuary as a framework to investigate environmental evolution and point towards clay grain coat potential.

• Recursive feature elimination (RFE) of geochemical elements found that the five most important elements are (in order of decreasing importance): Si, Pb, Al, Ca, and Mg. RFE also showed that some elements provide no information for the machine learning models or decrease overall model accuracy, and so have been removed from training data.

• Sediment geochemistry can be used to predict sub-depositional environment (Model 1) and estuarine zone (Model 2) in the Ravenglass Estuary using with an average overall accuracy of 57.90 % and 82.98 % respectively.

• The most accurate sub-depositional environments in geochemical models are mud flat (85.36 %) and mixed flat (78.62 %), likely due to differences in the abundance of mud-grade (<63 μm) sediment, linked with an increased proportion of clay minerals, and therefore varying geochemistry.

• Integrating geochemical and textural data to predict sub-depositional environment improves performance in all but one class, and increases overall model accuracy by 5.86 % (compared to texture-only models) and 15.7 % (compared to geochemistry-only models) to predict sub-depositional environment (Model 1). Integration of the geochemical and textural data also results in improvements of 3.23 % (compared to texture-only models) and 6.5 % (compared to geochemistry-only models) to predict estuarine zone (Model 2).

• The machine learning models created here to predict sub-depositional environment have been used to classify the palaeo-sub-depositional environment in three cores from a modern-day tidal flat in the Ravenglass Estuary. The ML output suggests an environmental transition occurred, possibly related to sea level change or spit development, which sheltered the central basin from wave action and allowed deposition of clay from suspension.

• In contrast to traditional core description and facies analysis, we are able to interpret the sub-depositional environment distribution in core with quantitative links to surface environments, providing a scheme for further analysis to understand grain coat potential.

• The mixed flat sub-depositional environment in the Ravenglass Estuary has been linked to the 'Goldilocks zone' of optimum potential for reservoir quality-enhancing grain coats at the surface and; this environment can be accurately predicted using a geochemistry-based machine learning model (up to 83.3 % accuracy).

• Using this study as an analogue, it is here suggested that employing the machine learning method outlined here on deeply buried clastic marginal marine systems to predict facies based on geochemistry and/or texture could serve to improve subsurface characterisation models of reservoirs. However, attention must be paid to understand how post-depositional diagenetic processes may have altered geochemical and textural properties since deposition which could affect classification models.

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

This work was undertaken as part of the Chlorite Consortium at the University of Liverpool, sponsored by BP, Equinor, and DNO. The authors gratefully acknowledge the comments and suggestions of an anonymous reviewer and Benjamin Brigaud that improved the paper.

References

Ainsworth, R.B., Vakarelov, B.K., Nanson, R.A., 2011. Dynamic spatial and temporal prediction of changes in depositional processes on clastic shorelines: toward improved subsurface uncertainty reduction and management. American Association of Petroleum Geologists Bulletin 95, 267–297.

Aitchison, J., 1982. The statistical analysis of compositional data. Journal of the Royal Statistical Society: Series B: Methodological 44, 139–160.

Ajdukiewicz, J.M., Lander, R.H., 2010. Sandstone reservoir quality prediction: The state of the artSandstone reservoir quality prediction: the state of the art. American Association of Petroleum Geologists Bulletin 94, 1083–1091.

Ajdukiewicz, J.M., Larese, R.E., 2012. How clay grain coats inhibit quartz cement and preserve porosity in deeply buried sandstones: observations and experiments. American Association of Petroleum Geologists Bulletin 96, 2091–2119.

Beard, D.C., Weyl, P.K., 1973. Influence of texture on porosity and permeability of unconsolidated sand. American Association of Petroleum Geologists Bulletin 57, 349–369.

Berner, R.A., 1980. Early diagenesis, a theoretical approach. Princeton University Press, Princeton.

Bjørlykke, K., Jahren, J., 2012. Open or closed geochemical systems during diagenesis in sedimentary basins: constraints on mass transfer during diagenesis and the prediction of porosity in sandstone and carbonate reservoirs. American Association of Petroleum Geologists Bulletin 96, 2193–2214.

Blott, S.J., Pye, K., 2001. GRADISTAT: a grain size distribution and statistics package for the analysis of unconsolidated sediments. Earth Surface Processes and Landforms 26, 1237–1248.

Boyd, R., Dalrymple, R.W., Zaitlin, B.A., 2006. Estuarine and incised-valley facies models. In: Posamentier, H.W., Walker, R.G. (Eds.), Facies Models Revisited. SEPM Society for Sedimentary Geology, Tulsa, Oklahoma, pp. 171–235.

Brockamp, O., Zuther, M., 2004. Changes in clay mineral content of tidal flat sediments resulting from dike construction along the Lower Saxony coast of the North Sea, Germany. Sedimentology 51, 591–600.

Cameron, D., Cooper, D., Johnson, E., Roberts, P., Cornwell, J., Bland, D., Nancarrow, P., 1993. Mineral Exploration in the Lower Palaeozoic Rocks of South-west Cumbria. Part 1, Regional Surveys.

Caracciolo, L., 2020. Sediment generation and sediment routing systems from a quantitative provenance analysis perspective: review, application and future development. Earth-Science Reviews 209.

Chan, M.A., 1985. Correlations of diagenesis with sedimentary facies in Eocene sandstones, western Oregon. Journal of Sedimentary Petrology 55, 322–333.

Chen, T., Guestrin, C., 2016. XGBoost. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Association for Computing Machinery, New York, New York, USA, pp. 785–794.

Chen, X.-w., Jeong, J.C., 2007. Enhanced recursive feature elimination. In: Sixth international conference on machine learning and applications (ICMLA 2007). IEEE, Cincinnati, Ohio, USA, pp. 429–435 (13-15 December 2007).

Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., Chen, K., Mitchell, R., Cano, I., Zhou, T., Li, M., Xie, J., Lin, M., Geng, Y., Li, Y., Jiaming, 2022. xgboost: Extreme Gradient Boosting.

Churchill, J.M., Poole, M.T., Skarpeid, S.S., Wakefield, M.I., 2017. Stratigraphic architecture of the Knarr Field, Norwegian North Sea: sedimentology and biostratigraphy of an evolving tide- to wave-dominated shoreline system. In: Hampson, G.J., Reynolds, A.D., Kostic, B., Wells, M.R. (Eds.), Sedimentology of Paralic Reservoirs: Recent Advances. Geological Society Special Publication, pp. 35–58.

Coleman, C.G., Grimoldi, E., Woollard, H., Holton, D., Shevelan, J., 2021. Developing 3D geological and hydrogeological models for the Low Level Waste Repository site, west Cumbria, UK. Quarterly Journal of Engineering Geology and Hydrogeology 54.

Dalrymple, R.W., Choi, K., 2007. Morphologic and facies trends through the fluvial–marine transition in tide-dominated depositional systems: a schematic framework for environmental and sequence-stratigraphic interpretation. Earth-Science Reviews 81, 135–174.

Dalrymple, R.W., Zaitlin, B.A., 1994. High-resolution sequence stratigraphy of a complex, incised valley succession, Cobequid Bay - Salmon River Estuary, Bay of Fundy, Canada. Sedimentology 41, 1069–1091.

Dalrymple, R.W., Zaitlin, B.A., Boyd, R., 1992. Estuarine facies models; conceptual basis and stratigraphic implications. Journal of Sedimentary Research 62, 1130–1146.

Deng, T., Xu, C., Lang, X., Doveton, J., 2021. Diagenetic facies classification in the Arbuckle Formation using deep neural networks. Mathematical Geosciences 53, 1491–1512.

Dowey, P.J., Hodgson, D.M., Worden, R.H., 2012. Pre-requisites, processes, and prediction of chlorite grain coatings in petroleum reservoirs: a review of subsurface examples. Marine and Petroleum Geology 32, 63–75.

Dubois, M.K., Bohling, G.C., Chakrabarti, S., 2007. Comparison of four approaches to a rock facies classification problem. Computers & Geosciences 33, 599–617.

Duller, R., Whittaker, A., Fedele, J., Whitchurch, A., Springett, J., Smithells, R., Fordyce, S., Allen, P., 2010. From grain size to tectonics. Journal of Geophysical Research - Earth Surface 115.

Ehrenberg, S., 1993. Preservation of anomalously high porosity in deeply buried sandstones by grain-coating chlorite: examples from the Norwegian continental shelf. American Association of Petroleum Geologists Bulletin 77, 1260–1286.

Fenies, H., Tastet, J.-P., 1998. Facies and architecture of an estuarine tidal bar (the Trompeloup bar, Gironde Estuary, SW France). Marine Geology 150, 149–169.

Folk, R.L., Ward, W.C., 1957. Brazos river bar. A study in the significance of grain size parameters. Journal of Sedimentary Petrology 27, 3–26.

Freedman, R., Herron, S., Anand, V., Herron, M., May, D., Rose, D., 2015. New method for determining mineralogy and matrix properties from elemental chemistry measured by gamma ray spectroscopy logging tools. SPE Reservoir Evaluation & Engineering 18, 599–608.

Garzanti, E., Andò, S., Vezzoli, G., 2009. Grain-size dependence of sediment composition and environmental bias in provenance studies. Earth and Planetary Science Letters 277, 422–432.

Gould, K., Pe-Piper, G., Piper, D.J.W., 2010. Relationship of diagenetic chlorite rims to depositional facies in Lower Cretaceous reservoir sandstones of the Scotian Basin. Sedimentology 57, 587–610.

Griffiths, J., Worden, R.H., Wooldridge, L.J., Utley, J.E.P., Duller, R.A., 2018. Detrital clay coats, clay minerals, and pyrite: a modern shallow-core analogue for ancient and deeply buried estuarine sandstones. Journal of Sedimentary Research 88, 1205–1237.

Griffiths, J., Worden, R.H., Wooldridge, L.J., Utley, J.E.P., Duller, R.A., 2019a. Compositional variation in modern estuarine sands: predicting major controls on sandstone reservoir quality. American Association of Petroleum Geologists Bulletin 103, 797–833.

Griffiths, J., Worden, R.H., Wooldridge, L.J., Utley, J.E.P., Duller, R.A., Edge, R.L., 2019b. Estuarine clay mineral distribution: modern analogue for ancient sandstone reservoir quality prediction. Sedimentology 66, 2011–2047.

Griffiths, J., Worden, R.H., Utley, J.E.P., Brostrøm, C., Martinius, A.W., Lawan, A.Y., Al-Hajri, A.I., 2021. Origin and distribution of grain-coating and pore-filling chlorite in deltaic sandstones for reservoir quality assessment. Marine and Petroleum Geology 134.

Guo, Y., Yang, S., Su, N., Li, C., Yin, P., Wang, Z., 2018. Revisiting the effects of hydrodynamic sorting and sedimentary recycling on chemical weathering indices. Geochimica et Cosmochimica Acta 227, 48–63.

Hansen, H.N., Lovstad, K., Muller, R., Jahren, J., 2017. Clay coating preserving high porosities in deeply buried intervals of the Sto Formation. Marine and Petroleum Geology 88, 648–658.

Hansen, H.N., Haile, B.G., Müller, R., Jahren, J., 2023. New direction for regional reservoir quality prediction using machine learning-example from the Stø Formation, SW Barents Sea, Norway. Journal of Petroleum Science and Engineering 220, 111149.

Heap, A.D., Bryce, S., Ryan, D.A., 2004. Facies evolution of Holocene estuaries and deltas: a large-sample statistical study from Australia. Sedimentary Geology 168, 1–17.

Herron, M.M., 1988. Geochemical classification of terrigenous sands and shales from core or log data. Journal of Sedimentary Research 58, 820–829.

Herron, M.M., Herron, S.L., 1990. Geological applications of geochemical well logging. Geological Applications of Wireline Logs 48, 165–175.

Hertzog, R., Colson, L., Seeman, B., O'Brien, M., Scott, H., McKeon, D., Wraight, P., Grau, J., Ellis, D., Schweitzer, J., 1989. Geochemical logging with spectrometry tools. SPE Formation Evaluation 4, 153–162.

Houghton, J.E., Nichols, T.E., Griffiths, J., Simon, N., Utley, J.E.P., Duller, R.A., Worden, R.H., 2023. Automated classification of estuarine sub-depositional environment using sediment texture. Journal of Geophysical Research - Earth Surface 128, e2022JF006891.

Kelly, M., Emptage, M., Mudge, S., Bradshaw, K., Hamilton-Taylor, J., 1991. The relationship between sediment and plutonium budgets in a small macrotidal estuary - Esk Estuary, Cumbria, UK. Journal of Environmental Radioactivity 13, 55–74.

Kuhn, M., Wickham, H., 2020. Tidymodels: a collection of packages for modeling and machine learning using tidyverse principlesTidymodels.

Lambiase, J.J., 1980. Hydraulic control of grain-size distributions in a macrotidal estuary. Sedimentology 27, 433–446.

Liu, M., Hu, S., Zhang, J., Zou, Y., 2023. Methods for identifying complex lithologies from log data based on machine learning. Unconventional Resources 3, 20–29.

Lloyd, J.M., Zong, Y., Fish, P., Innes, J.B., 2013. Holocene and Lateglacial relative sea-level change in north-west England: implications for glacial isostatic adjustment models. Journal of Quaternary Science 28, 59–70.

Løvstad, K., Hansen, H.N., Jahren, J., 2022. The porosity preserving effect of basin wide illitic coating in deeply buried sandstone intervals of the lower Jurassic Stø Formation, Barents Sea. Marine and Petroleum Geology 137, 105498.

Martinius, A., Ringrose, P.S., Brostrøm, C., Elfenbein, C., Næss, A., Ringas, J., 2005. Reservoir challenges of heterolithic tidal sandstone reservoirs in the Halten Terrace, mid-Norway. Petroleum Geoscience 11, 3–16.

McGhee, C.A., Muhammed, D.D., Simon, N., Acikalin, S., Utley, J.E.P., Griffiths, J., Wooldridge, L.M., Verhagen, I.T.E., van der Land, C., Worden, R.H., 2022. Stratigraphy and sedimentary evolution of a modern macro-tidal incised valley–an analogue for reservoir facies and architecture. Sedimentology 69, 696–723.

McLennan, S.M., 1993. Weathering and global denudation. The Journal of Geology 101, 295–303.

Merritt, J.W., Auton, C.A., 2000. An outline of the lithostratigraphy and depositional history of Quaternary deposits in the Sellafield district, west Cumbria. Proceedings of the Yorkshire Geological Society 53, 129–154.

Morad, S., Ketzer, J., De Ros, L.F., 2000. Spatial and temporal distribution of diagenetic alterations in siliciclastic rocks: implications for mass transfer in sedimentary basins. Sedimentology 47, 95–120.

Muhammed, D.D., Simon, N., Utley, J.E.P., Verhagen, I.T.E., Duller, R.A., Griffiths, J., Wooldridge, L.J., Worden, R.H., 2022. Geochemistry of sub-depositional environments in estuarine sediments: development of an approach to predict palaeo-environments from Holocene cores. Geosciences 12.

Nesbitt, H.W., Young, G.M., 1996. Petrogenesis of sediments in the absence of chemical weathering: effects of abrasion and sorting on bulk composition and mineralogy. Sedimentology 43, 341–358.

R Core Team, 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Roser, B., Korsch, R., 1986. Determination of tectonic setting of sandstone-mudstone suites using SiO2 content and K2O/Na2O ratio. The Journal of Geology 94, 635–650.

Saporetti, C.M., da Fonseca, L.G., Pereira, E., 2019. A lithology identification approach based on machine learning with evolutionary parameter tuning. IEEE Geoscience and Remote Sensing Letters 16, 1819–1823.

Simon, N., Worden, R.H., Muhammed, D.D., Utley, J.E.P., Verhagen, I.T.E., Griffiths, J., Wooldridge, L.J., 2021. Sediment textural characteristics of the Ravenglass Estuary; Development of a method to

predict palaeo sub-depositional environments from estuary core samples. Sedimentary Geology 418, 105906.

Smith, N.T., Merritt, J.W., Phillips, E.R., 2022. High-resolution 3D geological modelling of heterogeneity in poorly exposed glacial deposits using sedimentary and glaciotectonic architectural element analysis: a case example from Sellafield in west Cumbria, UK. Quarterly Journal of Engineering Geology and Hydrogeology 56.

Sokolova, M., Lapalme, G., 2009. A systematic analysis of performance measures for classification tasks. Information Processing & Management 45, 427–437.

Sun, Z., Jiang, B., Li, X., Li, J., Xiao, K., 2020. A data-driven approach for lithology identification based on parameter-optimized ensemble learning. Energies 13.

Taylor, T.R., Giles, M.R., Hathon, L.A., Diggs, T.N., Braunsdorf, N.R., Birbiglia, G.V., Kittridge, M.G., Macaulay, C.I., Espejo, I.S., 2010. Sandstone diagenesis and reservoir quality prediction: models, myths, and reality. American Association of Petroleum Geologists Bulletin 94, 1093–1132.

Tharwat, A., 2020. Classification assessment methods. Applied Computing and Informatics 17, 168–192.

Tolosana-Delgado, R., Talebi, H., Khodadadzadeh, M., Van den Boogaart, K., 2019. On machine learning algorithms and compositional data. In: Proceedings of the 8th International Workshop on Compositional Data Analysis. pp. 3–8 3-8 June. (Terrassa, Spain).

van den Boogaart, K.G., Tolosana-Delgado, R., 2021. compositions: Compositional Data AnalysisCompositions: Compositional Data Analysis.

Van Hoang, L., Clift, P.D., Schwab, A.M., Huuse, M., Nguyen, D.A., Zhen, S., 2010. Large-scale erosional response of SE Asia to monsoon evolution reconstructed from sedimentary records of the Song Hong-Yinggehai and Qiongdongnan basins, South China Sea. Geological Society, London, Special Publications 342, 219–244.

Virolle, M., Brigaud, B., Bourillot, R., Fenies, H., Portier, E., Duteil, T., Nouet, J., Patrier, P., Beaufort, D., 2019a. Detrital clay grain coats in estuarine clastic deposits: origin and spatial distribution within a modern sedimentary system, the Gironde Estuary (south-west France). Sedimentology 66, 859–894.

Virolle, M., Brigaud, B., Luby, S., Portier, E., Feines, H., Bourillot, R., Patrier, P., Beaufort, D., 2019b. Influence of sedimentation and detrital clay grain coats on chloritized sandstone reservoir qualities: insights from comparisons between ancient tidal heterolithic sandstones and a modern estuarine system. Marine and Petroleum Geology 107, 163–184.

Virolle, M., Féniès, H., Brigaud, B., Bourillot, R., Portier, E., Patrier, P., Beaufort, D., Jalon-Rojas, I., Derriennic, H., Miska, S., 2020. Facies associations, detrital clay grain coats and mineralogical characterization of the Gironde estuary tidal bars: a modern analogue for deeply buried estuarine sandstone reservoirs. Marine and Petroleum Geology 114.

Walderhaug, O., 1996. Kinetic modeling of quartz cementation and porosity loss in deeply buried sandstone reservoirs. American Association of Petroleum Geologists Bulletin 80, 731–745.

Wang, C., Pan, Y., Chen, J., Ouyang, Y., Rao, J., Jiang, Q., 2020. Indicator element selection and geochemical anomaly mapping using recursive feature elimination and random forest methods in the Jingdezhen region of Jiangxi Province, South China. Applied Geochemistry 122, 104760.

Wooldridge, L.J., Worden, R.H., Griffiths, J., Thompson, A., Chung, P., 2017a. Biofilm origin of clay-coated sand grains. Geology 45, 875–878.

Wooldridge, L.J., Worden, R.H., Griffiths, J., Utley, J.E.P., 2017b. Clay-coated sand grains in petroleum reservoirs: understanding their distribution via a modern analogue. Journal of Sedimentary Research 87, 338–352.

Wooldridge, L.J., Worden, R.H., Griffiths, J., Utley, J.E.P., 2018. The origin of clay-coated sand grains and sediment heterogeneity in tidal flats. Sedimentary Geology 373, 191–209.

Wooldridge, L.J., Worden, R.H., Griffiths, J., Utley, J.E.P., 2019a. Clay coat diversity in marginal marine sediments. Sedimentology 66, 1118–1138.

Wooldridge, L.J., Worden, R.H., Griffiths, J., Utley, J.E.P., 2019b. How to quantify clay-coat grain coverage in modern and ancient sediments. Journal of Sedimentary Research 89, 135–146.

Worden, R.H., Burley, S.D., 2003. Sandstone diagenesis: the evolution from sand to stone. In: Burley, S.D., Worden, R.H. (Eds.), Sandstone Diagenesis, Recent and Ancient. In: International Association of Sedimentologists Reprint Series. pp. 3–44.

Worden, R.H., Utley, J.E.P., 2022. Automated mineralogy (SEM-EDS) approach to sandstone reservoir quality and diagenesis. Frontiers in Earth Science 10, 794266.

Worden, R.H., Griffiths, J., Wooldridge, L.J., Utley, J.E.P., Lawan, A.Y., Muhammed, D.D., Simon, N., Armitage, P.J., 2020. Chlorite in sandstones. Earth-Science Reviews 204, 103105.

Xia, H.R., Perez, E.H., Dunn, T.L., 2020. The impact of grain-coating chlorite on the effective porosity of sandstones. Marine and Petroleum Geology 115, 104237.

Xie, Y., Zhu, C., Zhou, W., Li, Z., Liu, X., Tu, M., 2018. Evaluation of machine learning methods for formation lithology identification: a comparison of tuning processes and model performances. Journal of Petroleum Science and Engineering 160, 182–193.

Yasin, Q., Sohail, G.M., Khalid, P., Baklouti, S., Du, Q., 2021. Application of machine learning tool to predict the porosity of clastic depositional system, Indus Basin, Pakistan. Journal of Petroleum Science and Engineering 197, 107975.