

**Machine learning techniques applied to sediment core scanning
data in the framework of sedimentological and
paleoceanographical investigations**

Dissertation

A dissertation submitted to
Fachbereich 8, University of Bremen
and
Department of Geosciences, National Taiwan University

For the degree
Doctor of Natural Science (Dr. rer. Nat.) from the University of Bremen
and
Doctor for Philosophy (PhD) from National Taiwan University

Submitted by
M.Sc. An-Sheng Lee

June 2022

Referees and Committees

Prof. Dr. Bernd Zolitschka

Institute of Geography

University of Bremen, Germany

Prof. Dr. Sofia Ya Hsuan Liou

Department of Geosciences

National Taiwan University, Taiwan

Dr. Christian Ohlendorf

Institute of Geography

University of Bremen, Germany

Prof. Dr. Hsuan-Tien Lin

Department of Computer Science &

Information Engineering

National Taiwan University, Taiwan

Public defense: 16.08.2022

Acknowledgements

I am grateful to have my PhD career as an adventure. Along the path, I found plenty of opportunity and challenges. Sincere thanks to my supervisors Bernd and Sofia. With your mentoring, I could find ways to grab those opportunities and complete the challenges as my fuel of life. My vision has been extended from Taiwan to Europe and to the world after these years. In the meantime, without the help from Dirk, Christian, Hsuan-Tien, Jyh-jaan and Rafael, I would have been defeated by trivial details in my research. Special thanks to my buddies who made my world awesome and continuously having energy for the challenges: Stella, Britta, Ines, JC, Jose, Tenzin and Yang. May all the fun and joy be with you.

Danke schön.

This degree enriches my life my knowledge, but also sacrifices some parts of me to achieve. What can I say? Let's continue, shall we?

Abstract

Studying the mother Earth to understand and predict its system process facilitates the evolution of science and human society toward a progressive and sustainable stage. One of the key research materials is sediment archive, which records various short-and long-term historical information, such as climatic, biological, geomorphological and human activity variations. With the progress in geoscience and computer science, this thesis presents three interdisciplinary approaches to solve practical problems in sedimentological and paleoceanographical investigations.

The first study compiles different core scanning data (magnetic susceptibility, X-ray computed tomography, elemental profiles, digital photography) to characterize a priori classified sediment core sections recovered from the Wadden Sea region. The results confirm that the description of human-recognized sediment facies can be reproduced using the scanning data, which gives a promising hint to a further step: automatic sediment facies classification.

The second study increases the data scale by covering more core sections and sediment facies to develop a machine learning (ML) model that classifies sediments into facies by reading elemental profiles acquired from X-ray fluorescence (XRF) core scanning. A series of feature engineering and ML algorithms are tested to find the optimal solution. As a result, a simple but powerful model is proposed to simulate sedimentologists' observational behavior and have promising performance (78% accuracy), which is supported by a tailor-made evaluation involving sedimentary knowledge. Furthermore, the model can highlight critical sections of sediments requiring

sedimentologists' expertise. This provides an increased capability of classification without losing accuracy.

The third study focuses on obtaining cost-efficient high-resolution bulk chemistry measurement (CaCO_3 and total organic carbon) by quantifying XRF spectra using ML. This novel approach of using XRF spectra directly and enhanced regression power of ML eliminates manual bias and increases input information. Meanwhile, the previous limitation of data coverage is lifted by including multi-regional data (high latitude sectors of Pacific Ocean). The outcome is carefully evaluated using cross-validation, test set, and case study, with R^2 of CaCO_3 : 0.96 and TOC: 0.78 from the test set.

In conclusion, ML increases the capability of data analysis and automation. Sediment core scanning provides thorough but rapid measurement in high spatial resolution. This thesis offers generalizable methodological blueprints for integrating these two techniques to release the wealth of sedimentary information from the shackles of expensive and observer-dependent data in the past.

Zusammenfassung

Das Studium der physischen Geographie hilft Prozesse des Systems Erde zu verstehen und vorherzusagen und erleichtert die Entwicklung der Wissenschaft und der menschlichen Gesellschaft hin zu einer fortschrittlichen und nachhaltigen Welt. Eines der wichtigsten Forschungsmaterialien ist das Sedimentarchiv, das kurz- und langfristige Informationen über klimatische, biologische und geomorphologische Schwankungen sowie menschliche Aktivitäten aufzeichnet. Mit den Fortschritten in den Geowissenschaften und der Informatik stellt diese Arbeit drei interdisziplinäre Ansätze zur Lösung praktischer Probleme in sedimentologischen und paläozeanographischen Untersuchungen vor.

Die erste Studie stellt verschiedene Kernscandaten (magnetische Suszeptibilität, Röntgen-Computertomographie, Elementprofile, digitale Fotografie) zusammen, um bereits klassifizierte Sedimentkernabschnitte zu charakterisieren, die aus der Wattenmeerregion gewonnen wurden. Die Ergebnisse bestätigen, dass die vom Menschen durchgeführte Beschreibung der Sedimentfazies anhand der Scandaten reproduziert werden kann. Dies ermöglicht einen weiteren Schritt: die automatische Klassifizierung der Sedimentfazies.

Die zweite Studie enthält eine erhöhte Datenskala, indem sie mehr Kernabschnitte und Sedimentfazies abdeckt und entwickelt dadurch ein Modell für maschinelles Lernen (ML). Das Modell klassifiziert die Sedimente in Fazies, indem Elementprofile von Röntgenfluoreszenz (XRF)-Kernscanning gelesen werden. Eine Reihe von Feature-Engineering- und ML-Algorithmen werden getestet, um die optimale Lösung zu finden. Das Ergebnis ist ein einfaches, aber leistungsstarkes Modell, welches das Beobachtungsverhalten von Sedimentologen simuliert und eine Genauigkeit von 78 %

hat. Darüber hinaus wurden weitere Evaluierungen unter Einbeziehung von sedimentologischem Wissen durchgeführt, um die Verlässlichkeit des Modells zu gewährleisten. Darüber hinaus kann das Modell kritische Abschnitte von Sedimenten hervorheben, die das Fachwissen von Sedimentologen erfordern. Dies stellt eine erhöhte Klassifikationsfähigkeit bereit, ohne an Genauigkeit zu verlieren.

Die dritte Studie konzentriert sich auf die Erzielung einer kosteneffizienten hochauflösenden Bulk-Chemie-Messung (CaCO_3 und organischer Gesamtkohlenstoff) durch Quantifizierung von XRF-Spektren mit ML. Dieser neuartige Ansatz der direkten Verwendung von XRF-Spektren und der verbesserten Regressionsleistung von ML eliminiert manuelle Verzerrungen und erhöht die Eingabeinformationen. Inzwischen wird die frühere Datenbeschränkung durch Einbeziehung multiregionaler Daten (Sektoren des Pazifischen Ozeans in hohen Breitengraden) aufgehoben. Das Ergebnis wird mithilfe von Kreuzvalidierung, Testset und Fallstudie sorgfältig bewertet, wobei R^2 von CaCO_3 : 0,96 und TOC: 0,78 aus dem Testset ergibt.

Zusammenfassend lässt sich sagen, dass die ML-Technik das Leistungsvermögen der Datenanalyse und Automatisierung erhöht. Die Sedimentkern-Scanning-Technik ermöglicht gründlichere, aber schnelle Messungen mit hoher räumlicher Auflösung. Diese Dissertation bietet verallgemeinbare methodische Entwürfe für die Integration dieser beiden Techniken, um die Fülle an Sedimentinformationen von den Fesseln vergangener teurer und beobachterabhängiger Daten zu befreien.

摘要

解析地球系統的運作機制能夠促進我們在科學以及人類社會上的進步與永續，其中一項重要研究材料為記錄著短中長期氣候、地形、生物與人類活動訊號的沉積物紀錄。結合地球科學與電腦科學兩大領域的進展，本論文展示三項解決常見於地球系統中沉積學與古海洋學研究之問題的跨領域研究。第一項研究彙整多種採自德國瓦登海域之岩心掃描數據進行數據分析以解析不同沉積相之特性，結果顯示這些沉積物紀錄中的掃描數據確實於不同沉積相中具有不同之特徵，故從中我們得知沉積學家的觀察能夠再現於掃描數據。於是第二項研究「沉積相自動辨識」接續進行，所挑選之岩心及沉積相數量較第一項研究顯著提升，目標在於透過機器學習技術建立能藉由讀取 X 光螢光岩心掃描儀（以下簡稱 XRF）所產生之元素資料進行沉積物之沉積相辨識的模型。研究流程將一系列之特徵轉換與機器學習演算法進行組合測試並模擬出類似沉積學家觀察模式之資料轉換，透過融入沉積學原理特製化之模型能力檢驗，獲得一組最佳化具 78% 準確度之模型。除了具沉積相辨識能力，該模型亦可標記出模型較不具信心之沉積物區段以便予沉積學家投入其專業進行判斷，如此便可提升辨識量能且不影響整體品質。第三項研究則著重於產生高解析度之沉積物定量化學性質（碳酸鈣及總有機碳），本方法透過機器學習演算法直接對 XRF 所產出之原始光譜數據以及定量數據進行非線性迴歸建模，從此便可使用該模型對快速產生之高解析 XRF 數據進行量化求得碳酸鈣及總有機碳之分佈；此外，本研究納入太平洋高緯度多處之沉積物紀錄以突破過去常受限於單一井位或地區之應用性，其碳酸鈣及總有機碳量化準確度（ R^2 ）分別為 0.96 與 0.78。總結來說，機器學習技術能夠提升資料分析以及自動化之能力，而岩心掃描技術則提供快速且高解析度之多項測量，本論文提供結合這兩種技術之方法藍圖以期能夠突破過去沉積物研究常見之費工耗時與主觀影響，進而獲得更全面之沉積物紀錄資訊。

Table of contents

Acknowledgements	I
Abstract.....	II
Zusammenfassung	IV
摘要	VI
1. Introduction	1
1.1. Sediment archives.....	1
1.2. Primary techniques	3
1.2.1. Sediment core scanning.....	3
1.2.2. Machine learning	3
1.3. Objective.....	5
2. Study I: Trials of core scanning techniques.....	7
2.1. Introduction	10
2.2. Methods and materials.....	13
2.2.1. Coring	13
2.2.2. Lithology description.....	14
2.2.3. Down-core scanning techniques	14
2.3. Results	18
2.3.1. Reducing dimensionality of elemental intensities.....	18
2.3.2. Facies description	19
2.3.3. Collective results of scanning techniques.....	24
2.4. Discussion.....	27
2.4.1. Implications of principal component analysis	27
2.4.2. Facies characterisation.....	28
2.5. Limits and outlook.....	32
2.6. Conclusions	35
2.7. Acknowledgements	36
2.8. Supplementary materials	37
3. Study II: Automatic sediment facies classification	45
3.1. Introduction	47
3.2. Results	51
3.2.1. Contribution of feature engineering	51
3.2.2. Misclassification of the model.....	54
3.2.3. Highlighting critical segments.....	57
3.3. Discussion.....	59

3.4. Methods	61
3.4.1. Data acquisition	61
3.4.2. Feature engineering	61
3.4.3. Model building and evaluation	63
3.5. Acknowledgements	65
3.6. Supplementary materials	66
3.6.1. Material and methods	66
3.6.2. Grid searching results	69
3.6.3. Test set results	71
4. Study III: Bulk chemistry quantification	74
4.1. Introduction	76
4.1.1. Importance and limitations of bulk geochemistry	76
4.1.2. Advantages of XRF core scanning and its quantification	77
4.2. Results	81
4.2.1. Determination of optimal models	81
4.2.2. Evaluations of test set and case study	81
4.2.3. Outcome: quantified data of CaCO ₃ and TOC	84
4.3. Discussion	84
4.4. Guidelines for applications	86
4.5. Conclusions	87
4.6. Materials and methods	88
4.6.1. Sediment cores and bulk measurements	88
4.6.2. Setting of the Avaatech XRF scanning and data compilation	89
4.6.3. Pilot test	90
4.6.4. Model training and evaluation	91
4.7. Acknowledgments	93
4.8. Supplementary materials	94
4.8.1. Detailed grid search results	94
4.8.2. Info and exported data	96
5. Concluding remarks and outlook	98
5.1. Trials of core scanning techniques	98
5.2. Automatic sediment facies classification	98
5.3. Bulk chemistry quantification	99
6. References	101

1. Introduction

1.1. Sediment archives

Studying the mother Earth allows us to understand its function, predict its behavior and facilitate human society toward a progressive and sustainable stage. Numerous fields in geoscience have been developed to cope with this demand, such as geomorphology, tectonic, paleoclimate and paleoceanography. One of the essential studying materials for geoscientists is sediment archive, which preserves physical and chemical signals of the environment through the sedimentary history¹. For the scientific approach, these signals depict the climate, biology, and geomorphology variations²⁻⁴. With relevant chronology control, these variations are able to draw a storyline of the earth functioning system and history. For engineering application, the investigated signals provide vital information for civil construction and activities⁵.

In the vast applications of sediment archives, classification of sediment facies is the most fundamental step in constructing the geological background. The sediment is the outcome of the sedimentary process. Each depositional environment has a unique combination of sedimentary processes to form specific sediment types and structures, defined as sediment facies jointly¹. After classifying sediment facies, the paleoenvironments are reconstructed to realize the variation of climate or hydrology^{3,6}. For instance, Streif (2004)³ collected sediment archives in the southern North Sea basin and identified the changes of sediment facies to reconstruct coastal evolution during the Quaternary, which was driven by the sea-level rising since the last glacial maximum. This transgression generated complex interactions with the relief of the inundated geomorphology and the local hydrology. Studying this provides a basis for predicting the future development of the coastal zone.

Understanding sediment facies also gives necessary knowledge for exploring resources⁷ and constructing civil facilities⁸. Fujii et al. (2015)⁷ carried out the sediment analysis accompanied with geophysical approaches to reveal the offshore subsurface structures, which exhibited the basic information of methane hydrate reservoir characterization. Coughlan et al. (2020)⁸ mapped the Irish Economic Exclusion Zone within the Irish Sea by considering the bathymetry and sediment classification to assess the possible impact on the infrastructure of offshore wind farms. This gives proper constraints for stakeholders like developers, consultants and maritime spatial planning authorities.

After having sediment facies as a prior understanding, measuring bulk chemistry variation is another common sediment archives application. As mentioned beforehand, the sediment records chemical signal through the sedimentary history, which provides valuable proxies for paleoclimate and paleoceanography^{6,9}. For example, studying calcium carbonate (CaCO_3) of deep-sea pelagic sediments help us discover atmospheric CO_2 concentration in the past since CaCO_3 is one of the key proxies depicting the ocean's regulating process to CO_2 ¹⁰. Kropelin et al. (2008)⁶ presents a delicate paleoenvironmental record covering the past 6000 years within the Sahar, which pictures the dynamic terrestrial and aquatic ecosystem responding to climate-driven moisture-balance. It is achieved based on careful research on sedimentological and geochemical data coupled with biological indicators.

As a result, the sediment research involving sediment facies classification and bulk chemistry analysis is on a demanding evolution requiring continuous effort of improvement and innovation to fulfill our desire to study the Earth.

1.2. Primary techniques

1.2.1. Sediment core scanning

Various core scanning techniques have been on board to increase measuring resolution and reduce laboratory costs in the last decades. For measuring physical properties (e.g., density and magnetic susceptibility) in sediments, multi-sensor core logger ¹¹ is introduced. To achieve higher-resolution 3D observation of density, the X-ray computed tomography technique is adopted from the medical discipline ^{12, 13}. Digital imaging is also widely implemented to characterize sediments (e.g., indicating climate proxies and recognizing particles) ¹⁴. For measuring chemical properties in sediments, X-ray fluorescence (XRF) core scanning techniques are introduced to rapidly acquire elemental profiles in massive amounts of sediments, which are frequently used to interpret paleoclimate and anthropogenic activity ¹⁵⁻¹⁷. Most recently, hyperspectral imaging is capable to measure both chemical and physical properties to resolve various climatic and biochemical proxies by handling specific spectral bands ¹⁸⁻²⁰.

With these advances, geoscientists now have more possibilities to obtain near-continuous (μm -scale resolution) and non-destructive measurements covering both physical and chemical data from natural archives with less cost. Thus, the fineness of observation and data size have been promoted to an unprecedented scale.

1.2.2. Machine learning

Machine learning (ML) has significantly expanded over the past two decades, owing to improving computing strength and data acquisition capability. It is a combination of computer science and statistics, which has benefited extensively from biology to cosmology to social science ^{21, 22}.

ML intends to train computers using exemplary data instead of manually craft complex decision rules to solve a given problem ²³. ML techniques can be summarized into three primary

categories. The one which is broadly applied is supervised learning ²⁴. The algorithms falling into this category learn from the training data, which have pairs of inputs and desired outputs, and build a model to predict the desired outcomes with given inputs ²¹. For instance, Kuwatani et al. (2014) ²⁵ used a support vector machine (SVM) algorithm to learn the pairs of geochemical proxies and tsunami/non-tsunami deposit labels from sediments. They then use the developed model to discover tsunami deposits in other sediments by inputting the geochemical proxies. Besides SVM, there are other popular algorithms called decision trees, random forest, Bayesian classifiers, logistic regression, kernel machines, and neural networks (i.e., deep learning) ²⁴.

In the other hand, unsupervised learning algorithms learn from the data without the desired outputs (i.e., labels) ²⁶. Common signal decomposition algorithms, e.g., principal component analysis (PCA), factor analysis (FA), non-negative matrix factorization (NMF), manifold learning and autoencoders belong to this category. In addition, this method contains clustering algorithms, like k-Means, agglomerate clustering, and density-based spatial clustering of applications with noise (DBSCAN) ^{24, 26-29}. They are often utilized to assist scientists discover, understand and interpret their data. For example, Bulian et al. (2019) ³⁰ interpreted the climatic variation based on the PCA reduced data dimension of elemental intensities. Gebhardt et al. (2013) ³¹ applied k-Means clustering on well-logging and geochemical data to classify different glacial phases. Due to the nature of revealing latent information, the unsupervised learning algorithms are often employed to transform original data as a preprocessing step for facilitating consecutive supervised learning algorithm's performance ²¹.

The third category is reinforcement learning, which is built on the concepts from the control-theory literature. It relies on the training data indicating an action's correctness instead of the specific input-output pairs like supervised learning does. This method is mainly utilized in psychology, neuroscience, and policy optimization fields ³²⁻³⁴.

1.3. Objective

More and more works exploiting ML have been published in the geoscience community. Models for recognizing lithology facies in the petroleum industry have been developed using seismic reflection and well logging (e.g., gamma-ray curve and borehole image) data. The reservoir or organic-rich zones can be automatically classified³⁵⁻³⁷. Bolton et al. (2020)³⁸ proposed ML methods for the geochemical discrimination of tephra sources. Jacq et al. (2020)³⁹ classified sedimentary structures in lake sediments using hyperspectral imaging.

However, ML's strength in geoscience studies is limited in small data dimensions and utility scales based on those applications. For instance, the spatial resolution and coverage of observation are restricted. The target class amount is constrained to a few choices. Thus, the automation empowered by ML is still at an early stage in this field. We propose a hypothesis that the progress of automation can be advanced by including a more extensive database. As the input data, both geographical coverage of sampling and measuring resolution will be increased. The target class amount will then be enriched to improve the applicability. Furthermore, the generalization of the ML approaches, which is the vital assessment of automation, will be carefully evaluated to guarantee our advance rigorously.

This thesis will go through the problems happening in the framework of sedimentological and paleoceanographical investigations. Laborious disadvantages in the conventional methods of sediment facies classification and bulk chemistry measurement are long-lasting desired to be solved. A large geo-database supported by GEOPOLAR and Alfred-Wegener-Institut is ready. Novel approaches integrating (1) more measurements from the sediment archives (mainly provided by XRF core scanning) and (2) tailor-made ML architecture will be presented. XRF core scanning is expected to give a solution to acquire high resolution and rapid chemical measurements, which depict the variation of sediment mineral

composition, biological signal and grain size. Several feature-engineering attempts are anticipated as the analogs of human observation behaviors and the treatments of instrument bias to promote the performance of ML. In the end, they are designed to overcome the limitations and prove the hypothesis. Ultimately, it shall bring our capacity of exploring Earth to a new level.

2. Study I: Trials of core scanning techniques

Facies characterisation of sediments from the East Frisian Wadden Sea (Germany) - new insights from down-core scanning techniques

An-Sheng Lee^{1,2}, Dirk Enters^{3,1}, Jürgen Titschack^{4,5}, and Bernd Zolitschka¹

¹University of Bremen, Institute of Geography, Germany (E-Mail: alee@uni-bremen.de)

²National Taiwan University, Department of Geosciences and Research Center for Future Earth, Taipei, Taiwan

³Lower Saxony Institute for Historical Coastal Research, Wilhelmshaven, Germany

⁴University of Bremen, MARUM - Center for Marine Environmental Sciences, Germany

⁵Senckenberg am Meer, Marine Research Department, Wilhelmshaven, Germany

Published

Netherlands Journal of Geosciences, 100, e8, Article e8. <https://doi.org/10.1017/njg.2021.6>

The printed version equals the published version with editorial changes.

Author contributions

A. -S. L.: Conceptualization, Methodology, Investigation, Data curation, Writing – original draft.

D. E.: Conceptualization, Methodology, Writing – review and editing.

J. T.: Conceptualization, Methodology, Writing – review and editing.

B. Z.: Supervision, Resources, Funding acquisition, Writing – review and editing

Keywords: sediment facies, CT scanning, μ -XRF scanning, magnetic susceptibility scanning, digital image scanning, coastal geology

Abstract

Sediment facies provide fundamental information to interpret palaeoenvironments, climatic variation, archaeological aspects and natural resource potentials since they are summary products of depositional processes, environmental conditions and biological activities for a given time and location. The conventional method of facies discrimination relies on macroscopic and/or microscopic determination of sediment structures combined with basic physical, chemical and biological information. It is a qualitative measure, depending on observer-dependent sedimentological descriptions, which cannot be reanalysed readily by further studies. Quantitative laboratory measurements can overcome this disadvantage, but are in need of large sample numbers and/or high temporal resolution, and are time-, labour- and cost-intensive. In order to facilitate an observer-independent and efficient method of facies classification, our study evaluates the potential of combining four non-destructive down-core scanning techniques: magnetic susceptibility (MS), X-ray computed tomography (CT), X-ray fluorescence (XRF) and digital photography. These techniques were applied on selected sections of sediment cores recovered around the island of Norderney (East Frisian Wadden Sea, Germany). We process and integrate the acquired scanning measurements of XRF elemental intensities, represented by principal components, MS, CT density and lightness of eight sediment facies previously recognised by conventional facies analysis: moraine, eolian/fluvial, soil, peat, lagoonal, sand flat, channel fill and beach-foreshore. A novel type of density plot is introduced to visualise the digitised sediment information that allows an observer-independent differentiation of these facies types. Thus, the presented methodology provides the first step towards automated supervised facies classification with the potential to reproduce human assessments in a fully reproducible and quantitative manner.

2.1. Introduction

Sediments preserve specific characteristics for each depositional environment. As these alter through time, a succession of sediment facies is generated. Each of them is defined as the sum of a sedimentary unit's primary characteristics with possible diagenetic alterations¹. As a result, classifying sediment records into sediment facies is vital for studying palaeoenvironmental variability, climatic variation, archaeological aspects and natural resource potentials. For instance, lake-level fluctuations identified by sediment facies changes in the Dead Sea indicate climatic variation⁴⁰. Sometimes climatic variation is accompanied by adaptation of human settlements e.g., Behre (2004)⁴¹. In addition to palaeoenvironmental reconstruction, geochemical proxies measured in sediments are useful to depict climatic variation but need to be interpreted carefully including background knowledge of respective sediment facies^{42, 43}. Also the detection of oil and gas reservoirs relies on lithological facies, especially as their exploration is coming to an end³⁵. Thus, discrimination of sediment facies and stratigraphic units needs to be observer-independent and cost-efficient to cope with such an abundance of applications.

The conventional approach of facies discrimination is mainly based on classical core description. Macroscopically and/or microscopically determined sediment structures and colours combined with basic physical, chemical and biological information, such as the hydrochloric acid (HCl) test for carbonates, and smear slides for identifying minerals and organisms (e.g., diatoms and foraminifers) are used to distinguish different sediment facies^{30, 44, 45}. This description is a qualitative measure relying on the experience of scientists. To increase observer-independent measurements that can be effortlessly reanalysed or adopted by further studies, quantitative laboratory measurements such as grain size, bulk geochemical and faunistic analyses need to be carried out. However, these quantitative methods are usually time-

consuming, labour- and cost-intensive as well as challenging for achieving high-resolution measurements.

The development of rapid high-resolution and non-destructive down-core scanning techniques has improved the situation for physical and chemical sediment parameters recently. For instance, X-ray fluorescence (XRF) core scanners offer elemental profiles with up to 100 μm down-core resolution and digital images with a maximum resolution of less than 50 μm per pixel^{15, 16}. Additionally, computed tomography (CT) scanning provides 3D radiographic images with up to 40 μm per voxel (volumetric pixel; Geotek Ltd (2018)⁴⁶) and density-related values in Hounsfield units (HU)⁴⁷ after calibration. At a considerably lower spatial resolution of 4 mm, magnetic susceptibility (MS) profiles provide information about the content of iron-bearing (magnetic) minerals⁴⁸. Data acquisition for all these techniques can take place within hours for ca. 1 m long core section. Combining these scanning parameters provides a digital alternative for thoroughly describing sedimentary records (semi-)quantitatively. These data compare well with other studies and use less time and labour with reference to quantitative methods relying on individual samples.

Our study area is the shallow coast of northwestern Germany with its high variability along the North Sea Basin³. This region has undergone rapid changes since the sea-level low stand during the Last Glacial Maximum, when the coastline was shifted southward by >800 km. With the rising sea level^{49, 50}, the environment changed from terrestrial (glacial moraines, eolian, fluvial, peat) via intertidal to shallow marine^{3, 41}, where barrier islands, lagoons, tidal flats and channels as well as marsh environments were formed^{51, 52}. Consequently, many different facies are preserved in the sediment records, all of which provide an excellent opportunity for testing down-core scanning techniques for a (semi-)quantitative facies description.

Several studies investigated sediments from the Wadden Sea. Some used conventional geochemical methods, such as elemental analyses, and grain-size analysis of discrete samples,

to discuss sediment dynamics in salt marsh, lagoonal and tidal-flat environments⁵³⁻⁵⁵. Element ratios are often applied in these studies, such as the Ca/Sr and Zr/Al ratios, to characterise the sediment. Some studies integrate palaeontological analyses to discover more details with regard to climatic oscillations and depositional evolution^{30, 56-60}. Since peats play an essential role in this coastal region, some studies focus on its formation's chemical and biological processes e.g., Dellwig et al. (2002)⁶¹. Furthermore, several studies stress that anthropogenic influences like land reclamation, dredging, dike building, and harbour activities in recent times are causing a loss of fine-grained sediments^{53, 62-64}.

This study combines measurements from four scanning techniques (MS, CT, μ -XRF, digital photography) to digitally characterise different facies and compare results with qualitative lithological descriptions. It aims at evaluating the potential of high-resolution physical and chemical measurements provided by down-core scanning techniques as a means of reproducing sedimentological observations in a digital perspective. We want to test whether scanning parameters and facies determined by macroscopical observations differ and if this information can later be used for automatic facies discrimination.

2.2. Methods and materials

2.2.1. Coring

Our study area is located in the East Frisian Wadden Sea near the island of Norderney, Germany, at the southern margin of the North Sea shelf (Figure 2-1). More than 140 sediment cores (length: up to 6 m, diameter: 8 and 10 cm) were recovered in the framework of the interdisciplinary Wadden Sea Archive (WASA) project from tidal flats, channels and offshore (Figure 2-1). The cores were drilled using a Vibracorer (VKG-6), cut into ~1 m-long sections and stored at +4°C in the core repository of the GEOPOLAR lab, University of Bremen, Germany. For this study, 10 core sections up to 1.2 m long were selected (Figure 2-1; Table S 2-1 with basic information). These contain a representative selection of all major sediment facies retrieved up to the start of this study.

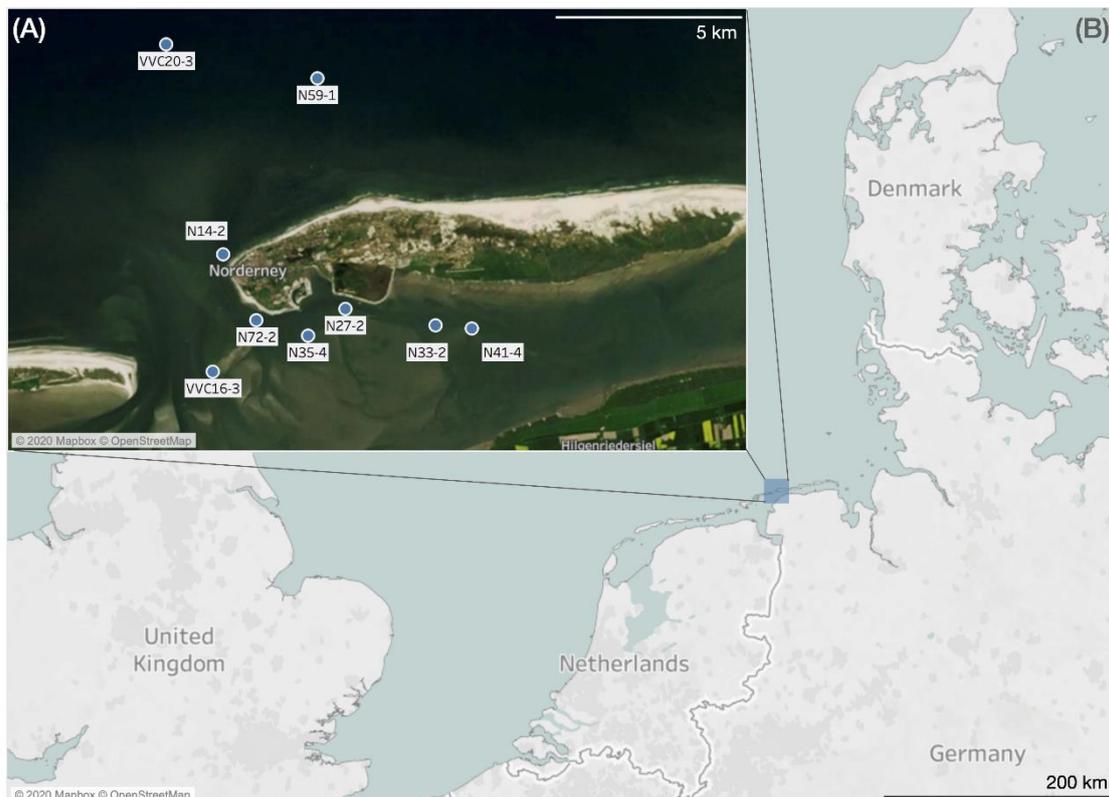


Figure 2-1. (A) Island of Norderney and sampling positions of analysed sediment core sections. Sections VVC16-3 and VVC16-4 belong to the same position. (B) Regional base map. The maps were captured from OpenStreetMap (<https://www.openstreetmap.org/copyright>).

2.2.2. Lithology description

The sediment cores were split lengthwise using a metal wire and sedimentary facies were described macroscopically. This description includes visible sediment structures, grain size, colour, carbonate content and biological remains. Occasionally, smear slides were made and selected sediment components were microscopically identified.

2.2.3. Down-core scanning techniques

The sediment surfaces of working halves were carefully cleaned with a razor blade prior to scanning. Core faces were covered by wrapping film during MS and CT scanning but not during μ -XRF scanning and digital photography. Measurements were taken exclusively in the centre of core to avoid possible coring artefacts (e.g., bending of layers) caused by the coring technique.

2.2.3.1. *Magnetic susceptibility*

The split-core surface's Magnetic susceptibility was scanned in 1 cm increments using a Bartington MS2F point sensor at the GEOPOLAR lab, University of Bremen⁴⁸. The sensitive volume of the probe is 15 mm (diameter) \times 6 mm (*depth*) Bartington Instruments Manual⁶⁵.

2.2.3.2. *X-ray density*

The split-core halves of the sediment sections were scanned with a Toshiba Aquilion 64 computed tomography (CT) with $0.35 \times 0.35 \times 0.5 \text{ mm}^3$ (x, y, z-direction) resolution at the hospital Klinikum Bremen-Mitte. The scanner used a voltage of 120 kV and a current of 600 mA as X-ray source parameters. Obtained X-ray topograms were reconstructed with Toshiba's patented helical cone-beam reconstruction technique and an overlapping resolution in z-direction of 0.3 mm. Each voxel provides X-ray attenuation at the given position. X-ray attenuation reflects the object density measured in Hounsfield units (HU), based on a linear

calibration of X-ray density with distilled water set to 0 HU and air set to -1000 HU ⁴⁷. X-ray attenuation is additionally influenced by the sediment's elemental composition ^{47, 66}.

The reconstructed three-dimensional X-ray attenuation (henceforth: CT density) data were analysed using the ZIB edition of the Amira software (version 2019.39; Stalling et al. (2005) ⁶⁷; <http://amira.zib.de>) to visualise and measure spatial variations of the internal composition. Within Amira and prior to all analyses, the core liners and about 2 mm of the outer core rims were deleted from the dataset. Sediment constituents, such as peat, rhizoliths, lithic clasts, shells and the sediment matrix were separated with a marker-based watershed algorithm, i.e., a *Watershed (Skeleton) Module*. Markers were set by threshold segmentation within the Segmentation Editor. CT-density value of the combined sediment matrix and peat constituents were determined by mean density per slice using the *Material Statistic Module*. Other constituents were excluded. To avoid coring artefacts, which are prominent especially at the outer rims of the cores, only values from the core centre (\varnothing : ~3.5 cm) were extracted.

2.2.3.3. *Elemental intensities and lightness (L^*)*

Elemental intensities and digital images were obtained directly from the split-core surfaces using a COX Itrax-XRF core scanner at the GEOPOLAR lab, University of Bremen. XRF measurements were taken with a 2 mm down-core scanning step size and a chromium X-ray tube and this setting: 30 kV voltage, 45 mA current and 5 s exposure time. The generated X-ray beam analysed a $20 \times 0.2 \text{ mm}^2$ wide rectangle with its long axis perpendicular to the core axis. Resulting XRF spectra were evaluated by the Q-Spec software (version 2015; COX Analytical Systems) to calculate elemental peak areas as total counts. Twelve elements (Si, S, Cl, K, Ca, Ti, Fe, Br, Rb, Sr, Zr, Ba) were selected based on the following criteria: 1) the fluorescence signal of the analytes must not be severely affected by the X-ray tube's primary

radiation and has maximum counts >500, and 2) median counts are >100 or coefficients of variance are >2.

We applied a centred log-ratio (clr) transformation on selected element counts to eliminate down-core bias caused by X-ray tube ageing, physical sediment variation, and the closed-sum effect⁶⁸⁻⁷⁰. Subsequently, a principal component analysis (PCA) was performed to reduce the data dimension with a minimal loss of variance and highlight major elements' behaviour^{71, 72}. However, to eliminate sample-size bias and compensate for differences in variance between elements during PCA calculation, 1) the number of data points belonging to each facies needs to be equal, and 2) each element profile needs to be standardised to unit variance and zero mean. Therefore, 500 data points were randomly selected without replacement in each facies that has >500 data points. For facies with fewer data points, 500 data points were randomly selected with replacement. The original data points for each facies are listed in Table S 2-2. We used the Python package scikit-learn⁷³ to train the PCA model with this homogenized and standardised data. In the model, the dimensions of elemental intensities were reduced to several principal components (PCs) based on the elbow concept (Figure S 2-1) and the preference for components having more than 10% of the total variance⁷⁴. Subsequently, the model is used to reduce the dimensions on the standardised data without prior homogenisation.

The optical-line camera is equipped with a light-sensitive 2048 pixel CMOS (complementary metal-oxide-semiconductor) device that generates high-resolution digital photographs ($47 \mu\text{m}/\text{pixel}$) of the sediment surface¹⁶. These digital images are in 8-bit RGB colour space designed for television and monitors. The colour space was transformed into CIE (Commission internationale de l'Éclairage) L*a*b* space, which is suited to depict colour variations closer to human perception and commonly accepted by sediment colour studies^{75, 76}. The Python package scikit-image⁷⁷ conducted the transformation with the most commonly

used settings (illuminant: D65, observer: 2) referring the pure white as seen by natural daylight¹⁴. L* variations were determined down-core by calculating the average L* value for each depth. All values were extracted from the central area (width: ~3.3 cm) to avoid coring artefacts at the outer core rim.

We developed a new type of plot (density plot) in order to combine all scanning results for making an intuitive comparison between facies. Density plots illustrate the probability density functions of our scanning results. This density function is a derivative of a histogram calculated by the Gaussian kernel density estimator (KDE). The concept of data distribution in a standardised and smoothed perspective, i.e., the unit probability cf.⁷⁸, is demonstrated in this plot type. Sample sizes and histograms of measurements for each facies are provided in Table S 2-2 and Figure S 2-4 - Figure S 2-7. All computations and visualisations were conducted using the SciPy ecosystem in Python^{73, 79-83}.

2.3. Results

2.3.1. Reducing dimensionality of elemental intensities

Three reduced dimensions (PCs) in the PCA model, preserving 80.3% of the total variance (Figure 2-2), were selected. The scree plot of the explained variance ratio for each PC is provided in Figure S 2-1. PC1 (47.0% variance explained) shows a negative correlation to the elements Si, K, Ti, Rb, Zr while it is positively correlated to the elements Br, S, Cl, and Fe. PC2 (19.6% variance explained) has a conspicuous co-variation for Ca and Sr. Besides, Fe and Ti have an opposite correlation to Ca and Sr. PC3 has a distinctively negative correlation for Ba.

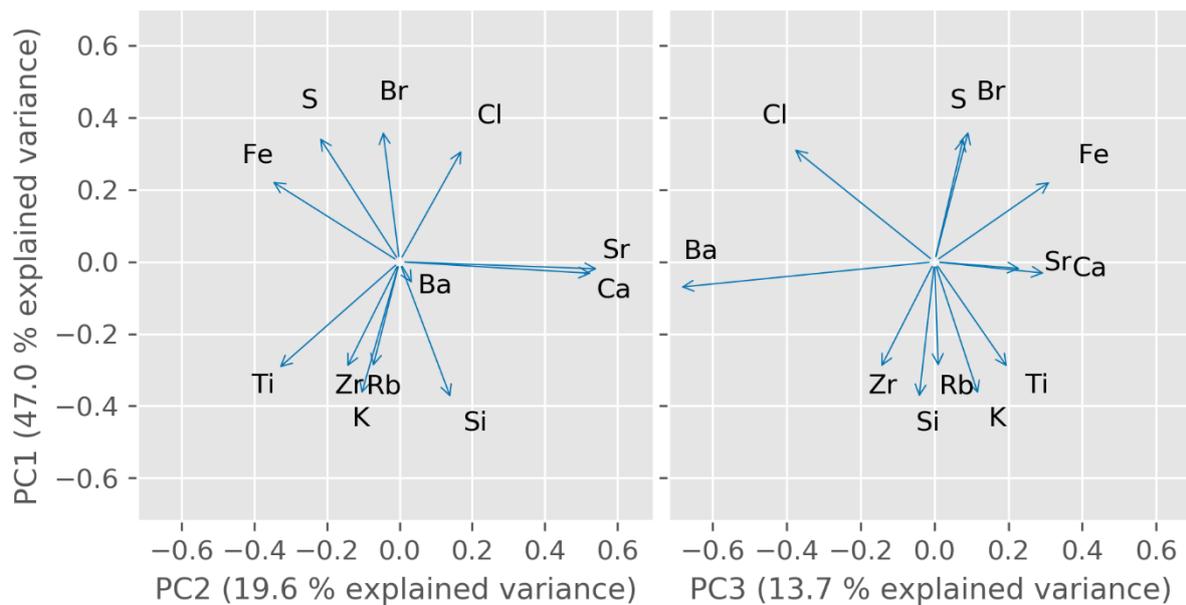


Figure 2-2. PCA loadings from homogenized elemental intensities. The first three principal components together preserve 80.3% of the total variance.

2.3.2. Facies description

There are eight facies (moraine, eolian/fluvial, soil, peat, lagoonal, sand flat, channel fill, and beach-foreshore) recognised based on lithology descriptions in the selected sections. The descriptions, according to macroscopic lithological observations accompanied by CT visualisation, are listed below. It should be noted that this observer-dependent facies classification is likely not directly comparable to a classification based on the integration of further (quantitative) measurements, e.g., palaeontological and grain-size data, and provides a rather general facies discrimination. Close-up photographs of the most important facies found in the WASA project are shown by Capperucci et al. (2022)⁸⁴.

Moraine sediments are characterised by compacted and unsorted sediments (Figure 2-3A). Their grain size varies from gravel to silt. Lithic clasts have angular shapes. The sediment has a greyish colour and lacks stratification. The carbonate content is usually low, but whitish carbonate clasts cause a strong reaction with HCl (10%).

Eolian/fluvial sediments are composed of well-sorted fine sand that lacks sedimentary structures and microfossils. Occasional stratifications and rhizoliths are observed (Figure 2-4). The carbonate content is very low. Further differentiation between eolian and fluvial sediments need further analyses, such as grain-size measurements.

Soils are composed of brown silt to fine sand with a high amount of organic matter and mark the end of the transition from eolian/fluvial sediments to peat with an unconformity at the top. The high organic matter content starts to decrease rapidly down-core (Figure 2-4). Also, multiple rhizoliths are visible in the 3D reconstruction (Figure 2-4d) originating from this horizon.

Peat contains dark brown to black deposits consisting mainly of organic matter. Plant fragments are the dominating components (Figure 2-4). Minerogenic components are rarely observed.

Finely bedded greyish silts and intercalated fine sands are observed in lagoonal facies (Figure 2-4), containing reworked peat and plant debris, often *Phragmites* (reed) fragments or rhizomes. The definition of this facies refers to sediments deposited in protected, quiet and shallow brackish environments, possibly behind barrier islands or levees^{3, 45}. Salt marsh sediments may share similar lithological characteristics, but palaeontological data are lacking so far for better discrimination.

Sand flat deposits consist of carbonate containing, greyish fine sands intercalated by shell fragments and occasional mud layers (Figure 2-5A). The macroscopically described grain-size varies from fine to medium sand.

Channel fill deposits comprise finely interlayered grey sands and dark grey muds, often showing flaser or lenticular beddings (Figure 2-3B). Shell fragments and bioturbation features are rare.

Beach-foreshore deposits consist of fine to coarse sand and layers of brownish shell fragments (Figure 2-5B). Their colour ranges from grey to dark grey. Few bedding planes have been recognised.

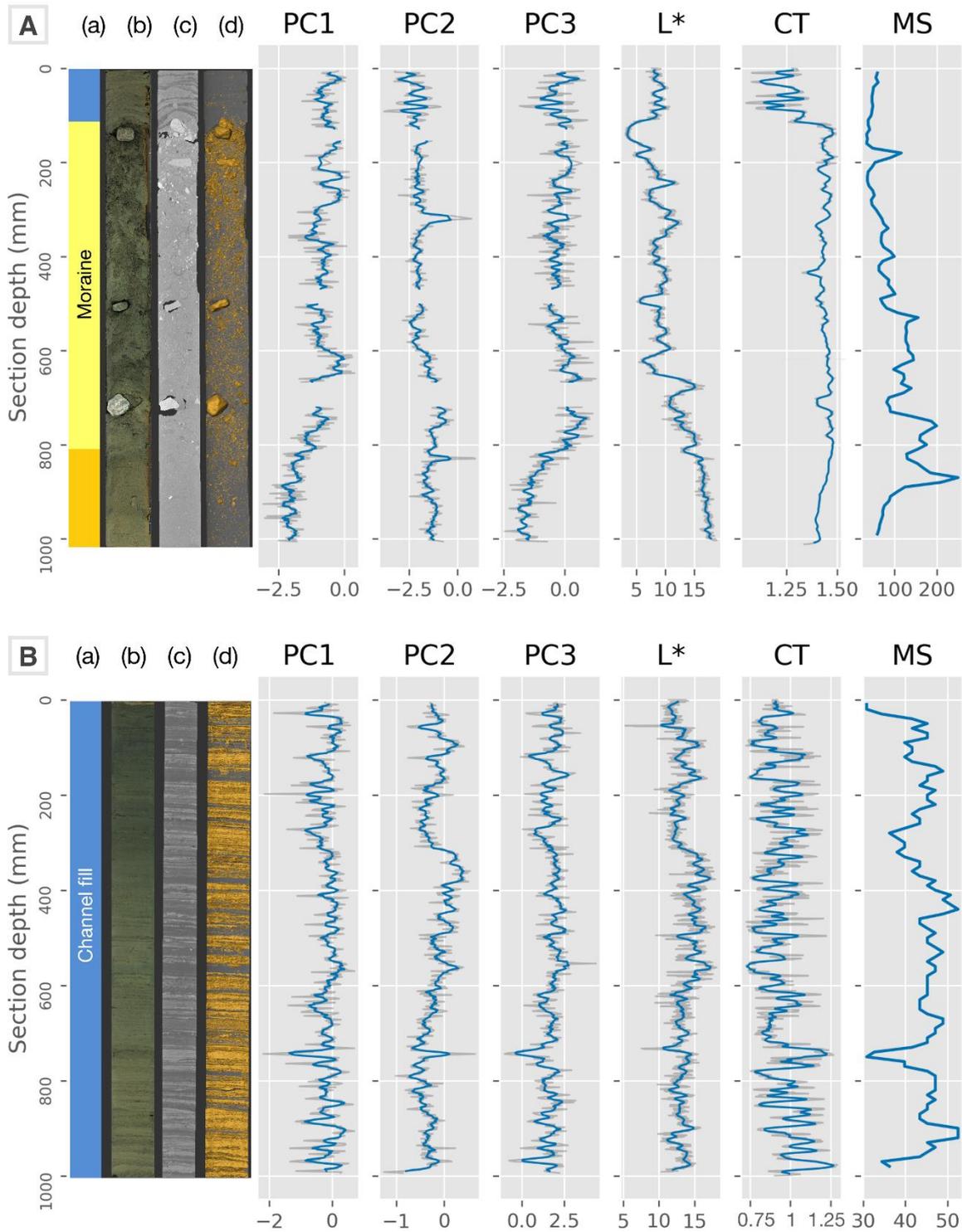


Figure 2-3. Cores N14-2 (A) and N33-2 (B): facies labels (a); digital photograph (b); radiograph (c); and sediment constituents (golden: high-density sediments and lithic clasts) separated using CT density (d). Sediment constituents (d) reveal intercalated layer structures and bioturbation. Scores of PC1–PC3 (dimensionless) from elemental intensities, lightness (L^*), ranging from 0 (black) to 100 (white), CT density (CT in 103 HU) and magnetic susceptibility (MS in 10^{-6} SI). Blue lines represent an 8 mm moving average, while grey lines represent raw data. No moving average is applied to MS.

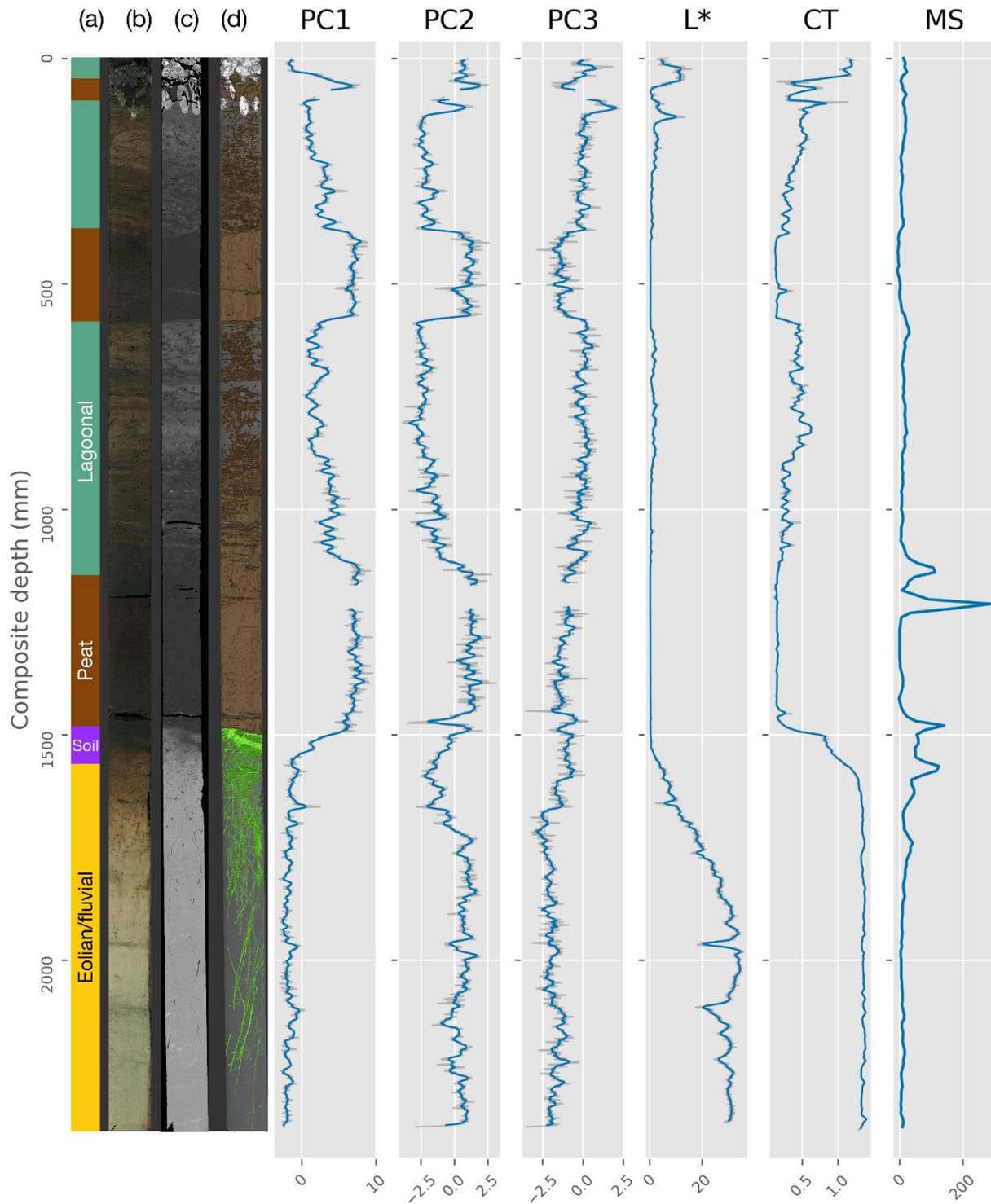


Figure 2-4. Core VVC16-3+4: facies labels (a); digital photograph (b); radiograph (c); and sediment constituents (brown: peat; green: rhizoliths; white: shell fragments) separated using CT density (d). Sediment constituents (d) reveal some hidden structures of the core, such as the deeply penetrating rhizoliths. Scores of PC1–PC3 (dimensionless) from elemental intensities, lightness (L^*), ranging from 0 (black) to 100 (white), CT density (CT in 103 HU) and magnetic susceptibility (MS in 10^{-6} SI). Blue lines represent an 8 mm moving average, while grey lines represent raw data. No moving average is applied to MS.

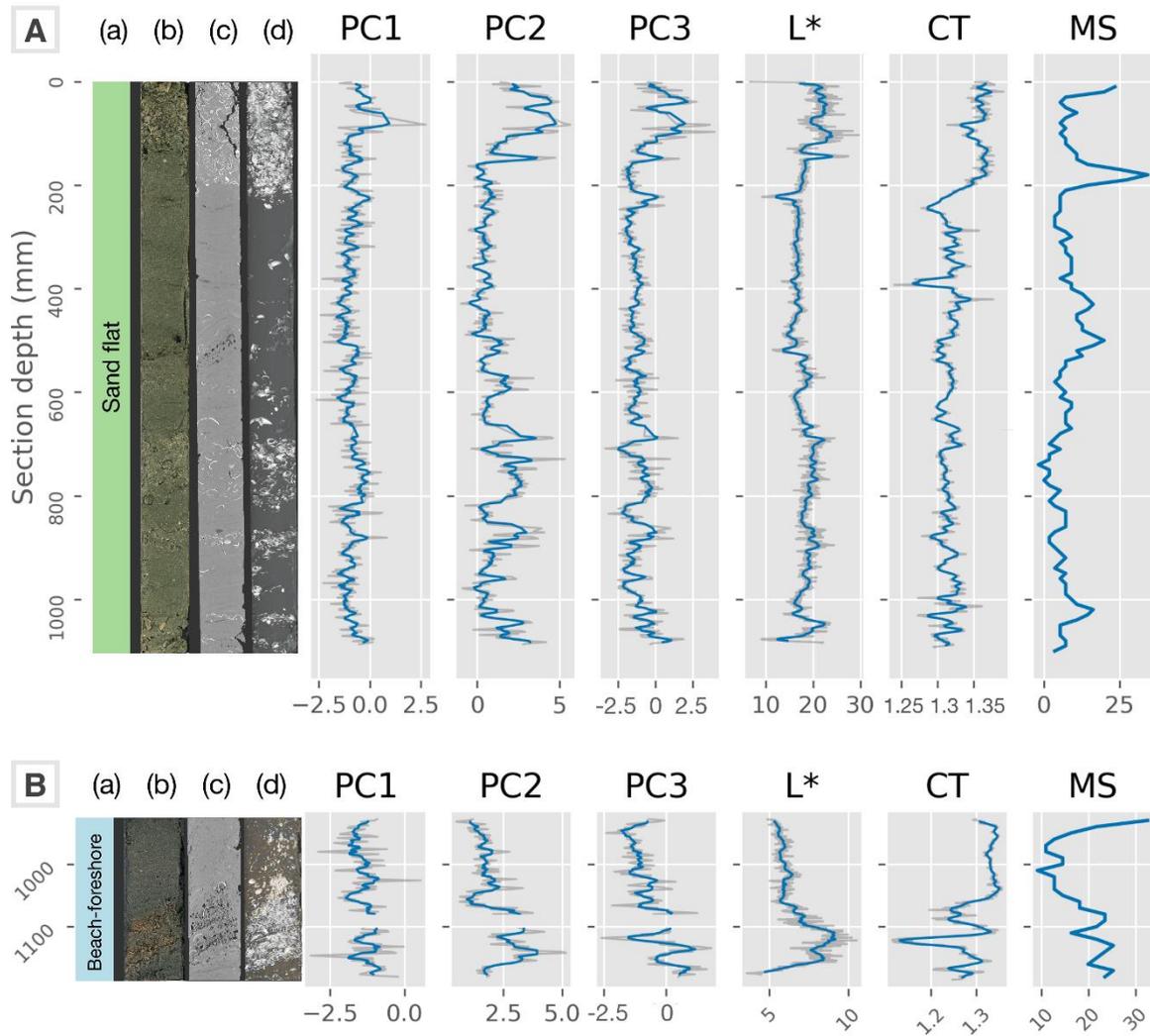


Figure 2-5. Cores N35-4 (A) and VVC20-3 (B): facies labels (a), digital photograph (b), radiograph (c) and sediment constituents (white: shell fragments; golden: high-density sediments and lithic clasts) separated using CT density (d). Scores of PC1–PC3 (dimensionless) from elemental intensities, lightness (L^*), ranging from 0 (black) to 100 (white), CT density (CT in 103HU), and magnetic susceptibility (MS in 10^{-6} SI). Blue lines represent an 8 mm moving average, while grey lines represent raw data. No moving average is applied to MS.

2.3.3. Collective results of scanning techniques

There are two types of visualisation for scanning measurements of the previously recognised facies. First, the measurements (PC1-3 scores, L*, CT density, and MS) are plotted on a depth scale along with facies labels, digital photographs, radiographs, and sediment constituents separated by CT density (Figure 2-3 - Figure 2-5). Variations of these measurements denote the spatial distribution of physical and chemical characteristics within and among facies. A steep increase in CT density marks the transition from moraine sediments to channel fill (Figure 2-3A), while the intersections of mud and sand in the channel fill facies itself is denoted by frequent wiggles (Figure 2-3B). The PC scores express the changes in peat sections (Figure 2-4), while L* and MS show only weak variations. The change from eolian/fluvial sediments to soil is documented by a gradual colour transition (L*), but rather sharp in CT density. Secondly, the measurements' data distributions are plotted for each facies (density plots: Figure 2-6 and Figure 2-7). The PCs scores range from -5 to 10, while L* varies from 0 to 40. The values of MS are spanning from 0 to 300×10^{-6} SI. CT-density values vary from 300 to 1700 HU. The shape of the data distributions illustrates the characteristics of each facies. Multiple peaks in a single analyte indicate mixed properties, whereas a single and narrow peak represents a uniform property. Each facies expresses a distinct behaviour based on the shapes and locations of distributions. Detailed characterisations are discussed below.

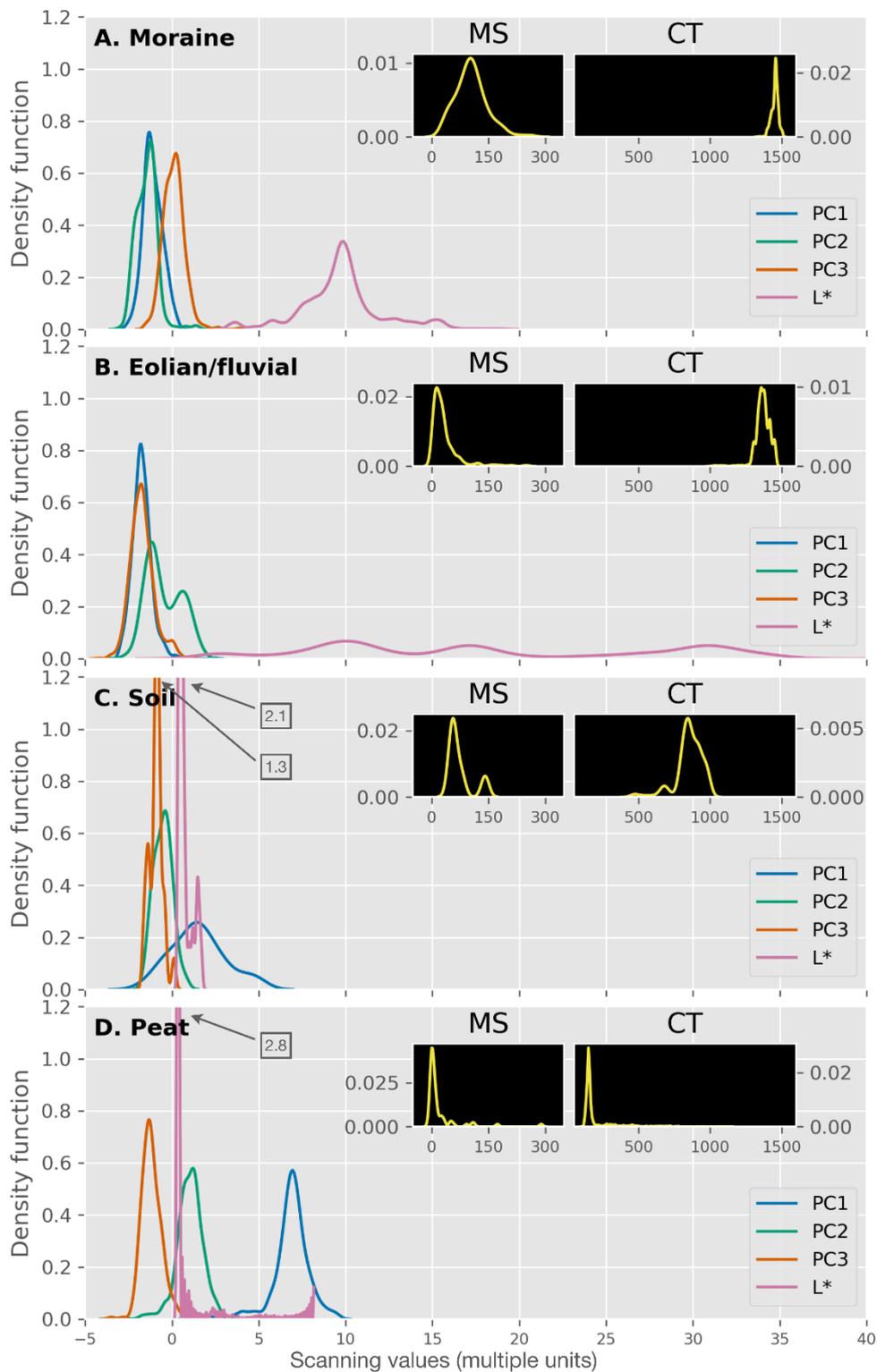


Figure 2-6. Density plots of down-core scanning results from terrestrial sediment facies (A: moraine; B: eolian/fluvial; C: soil; D: peat) showing the scores of PC1, PC2, PC3 and lightness (L*). The extreme peak values in density function are noted in text boxes connected with arrows. As inserts, the density plots for magnetic susceptibility (MS) and CT density (CT) are provided. Each curve represents a smoothed and standardised density function of each result. The area below stands for the probability, which sums up to 1. The x-axis is a pure number scale for placing different down-core scanning results on their own unit. For reasons of comparison, the axes are fixed to a certain range. For the raw histograms, which have independent axis ranges, please refer to Figure S 2-4 - Figure S 2-7.

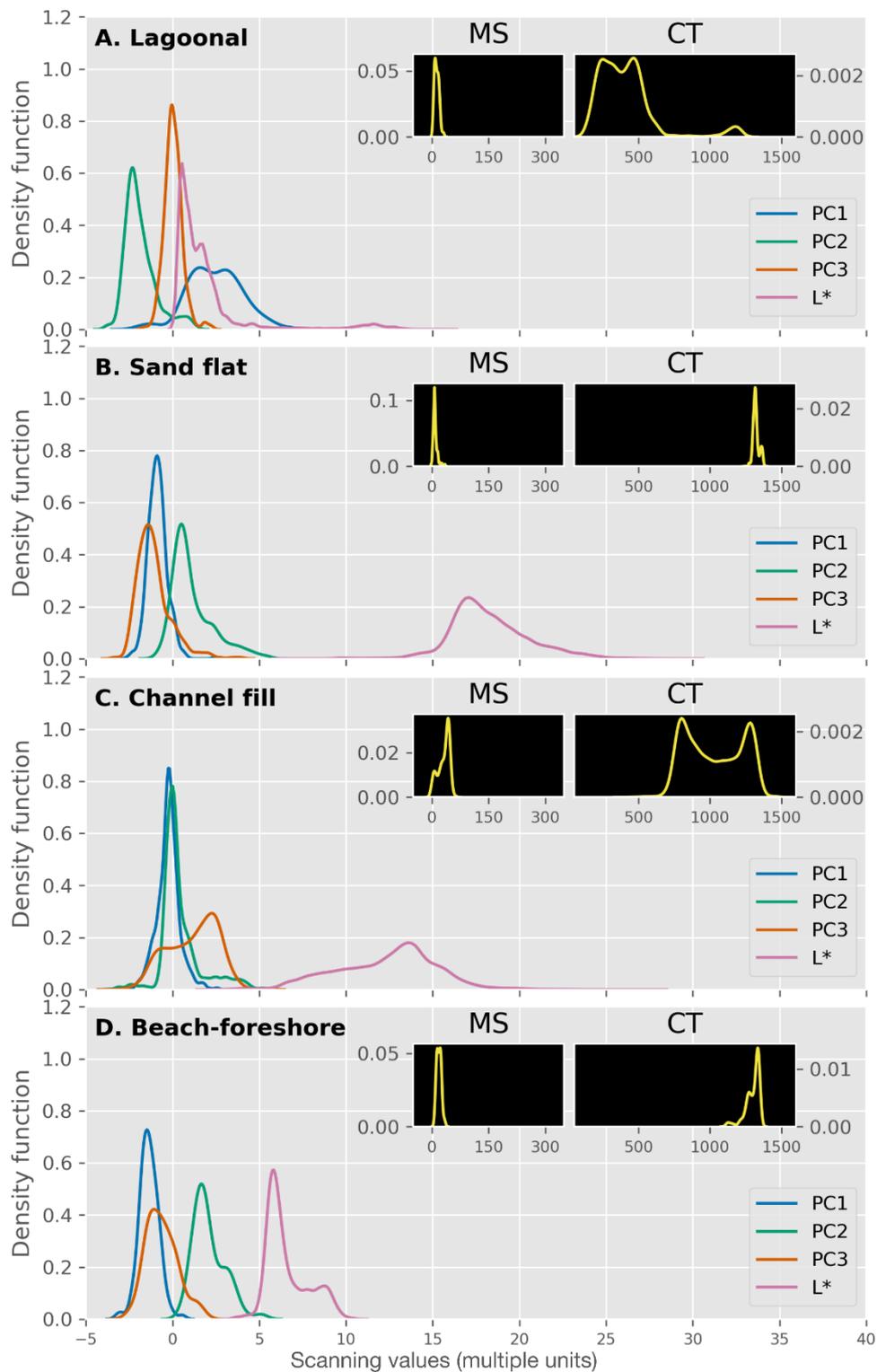


Figure 2-7. Density plots of down-core scanning results from marine sediment facies (A: lagoonal; B: sand flat; C: channel fill; D: beach-foreshore) showing the scores of PC1, PC2, PC3 and lightness (L^*). As inserts, the density plots for magnetic susceptibility (MS) and CT density (CT) are provided. Each curve represents a smoothed and standardised density function of each result. The area below stands for the probability, which sums up to 1. The x-axis is a pure number scale for placing different down-core scanning results on their own unit. For reasons of comparison, the axes are fixed to a certain range. For the raw histograms, which have independent axis ranges, please refer to Figure S 2-4 - Figure S 2-7.

2.4. Discussion

2.4.1. Implications of principal component analysis

PC1 shows a negative correlation to the elements related to rock-forming minerals (Si, K, Ti, Rb, Zr) (Figure 2-2) ^{43, 85}. It also has a positive correlation to Br, S, Cl, and Fe, indicating the existence of organic matter and pyrite formation ⁸⁶. High organic matter content in sediments produces a reducing environment for pyrite to form when accompanied by Fe available from the pore water ⁴³. In addition, the organic matter absorbs Br and Cl due to their high affinity ^{87, 88}. This interpretation is supported by the corresponding results of TOC and Br analyses from an accompanying study ³⁰. Without conventional grain-size measurements, the usefulness of a common grain-size proxy, e.g., Zr/Rb ⁸⁹, is limited in this study. As a result, PC1 is regarded as an index for distinguishing organic (positive) from lithogenic sediment components (negative). PC2 has a conspicuous co-variation for Ca and Sr, which links to biogenic carbonates' presence ⁴³. In addition, Fe and Ti, presenting in the clay fraction and heavy minerals ^{54, 55}, correlate negatively with Ca and Sr. Thus, PC2 is interpreted as enrichment of carbonates, which relates to marine influences ⁴³. PC3 has a distinctively negative correlation for Ba, an element often used as palaeoproductivity proxy ^{90, 91} in deep-sea environments. However, this interpretation cannot be transferred to coastal sediments with water depths of <100 m. Besides, it would conflict with PC1 as an expression of organic matter. Thus, the environmental condition causing PC3 (Ba) fluctuations remain unclear. So far, it is essential to note that the PCA depicts only the general covariance between elements of sediment records as a whole. Considering our study's contrasting depositional environments, a consistent interpretation of a specific element (or PC) across the whole dataset may not be feasible.

2.4.2. Facies characterisation

2.4.2.1. *Moraine sediments*

The three PCs' scores distribute in symmetrical and unimodal shapes centred at -1.5 and 0 (Figure 2-6A). PC1 and PC2 indicate a lithogenic and terrigenous nature of the sediments. L* supports the greyish colour with a slightly left-skewed distribution centred at a high value of 10. The high values and unimodal shapes of MS and CT density in Figure 2-6A indicate a high density and more iron-bearing (magnetic) minerals, which corresponds to the interpretation of PC1 and PC2. The lack of bimodal distributions due to the gravel-silt grain-size variance is because: 1) the elemental intensities at gravel sections were mostly excluded due to the low data quality for XRF core scanning, 2) the colours of gravels are not consistent, 3) the CT densities of gravels were excluded, 4) the appearance of gravels is relatively rare (Figure 2-3A). We only present one section for this facies in Figure 2-3A. For the other section containing moraine deposits, please refer to Figure S 2-2A.

2.4.2.2. *Eolian/fluviol sediments*

The PC1 score (Figure 2-6B) is distributed in a unimodal and symmetrical shape and centred at -2, suggesting a lack of organic matter. In contrast to the moraine facies, the PC3 score shifts to negative values and overlaps with PC1. For PC2, scores show a bimodal shape in distribution, located at opposite sides of zero. This may indicate an incomplete carbonate removal or unidentified/unrecognized Eemian marine sediments. Such sediments were found in other cores obtained within the WASA project⁹². The wide range of L* (Figure 2-6B), which contains three minor peaks, sets up a unique feature of this facies. Two of the peaks have locations coinciding with the unimodal peaks for moraine (10) and sand flat facies (17), supporting the multiple sand sources for this facies. The reworking of Saalian till by eolian processes mentioned by Schüttenhelm and Laban (2005)⁵⁹ could be one reason. The high CT density and low MS values (close to 0, Figure 2-6B) imply a predominantly lithogenic and

paramagnetic (i.e., quartz-rich) sediment composition. We only present two sections with this facies (Figure 2-3 - Figure 2-4). For the other section, please refer to Figure S 2-2A.

2.4.2.3. Soil

PC1 and PC2 scores reveal leaching evidence of the soil (Figure 2-6C). The increasing PC1 score and the slightly decreasing PC2 score compared to the eolian/fluvial facies imply accumulation of Fe and organic matter as well as the removal of carbonates due to leaching processes¹. The peak in MS (Figure 2-4) at around 1500 mm composite sediment depth indicates the presence of Fe-bearing minerals. The bimodal shape of L* (Figure 2-6C) with peaks located at 0.5 and 1.5 supports the darker colours. From the bimodal shape of MS and L* (Figure 2-6C), this weathered material has distinct differences in colour and Fe content compared to less weathered sediment. However, due to the small sample size of MS measurements for this facies, (Table S 2-2) the MS data are not robust. The relatively symmetrical, unimodal and wide shape of PC1 (Figure 2-6C) indicates the balancing mixture of lithogenic sediments, Fe-minerals and organic matter. The median value of CT density comparing to the eolian/fluvial and peat facies in Fig. 6C, and the variations of L* and CT density in Figure 2-4 indicate an intermediate stage between eolian/fluvial sediments and peat.

2.4.2.4. Peat

The lithological descriptions of peat are corroborated by scanning results (Figure 2-6D). The unimodal, symmetrical, and narrow shape of the PC1 score, located at the highest positive value (7) compared to the other facies types, implies a high organic matter content. The total organic carbon content of up to 38 % in peat layers, measured for a neighbouring core, supports this observation³⁰. This agrees with rare lithic components since MS and CT-density values remain close to zero (Figure 2-6D). Together with a (probably post-depositional) marine influence, revealed by slightly higher PC2 (ca. 1.5), an enhanced formation of pyrite, indicated by PC1, seems to describe the coincidence of coastal peat growth, seawater sulfate influence,

and microbial reduction^{61,93}. This phenomenon is supported by the intercalation of this facies with lagoonal sediments (brackish environment). Due to the homogeneously dark colour, values of L* also concentrate close to zero.

2.4.2.5. Lagoonal sediments

The multiple peaks found in the scanning data (PC1 score, CT density and L*) (Figure 2-7A) support the mixture of materials typical for such deposits¹. MS cannot detect this characteristic probably because its scanning step size is larger than the predominant thickness of the beddings or the difference in Fe-mineral content of this facies is indistinct. If the cause of indistinct Fe-mineral difference can be excluded, it would suggest a predominant bedding thickness smaller than the scanning step size of MS, i.e., 1 cm. The two peaks in the high PC1 score (2 and 4, Figure 2-7A) present a mixture of peat debris/organic material and fine-grained minerogenic matter. The three peaks found for CT density may indicate peat debris, lagoonal mud and few sand grains (from left to right). The bimodal shape of L* backs up these colour changes (Figure 2-7A). The unimodal and narrow shape of the PC2 score with a low value of -3 implies an absence of carbonates since they are dissolved by the acidic environment caused by the degradation of organic matter. This corresponds to the CaCO₃ undersaturation of the water mentioned by Alve and Murray (1995)⁹⁴. Kolditz et al. (2012)⁵⁵ also mention the carbonate-free characteristics of lagoonal sediments for the island of Langeoog. Furthermore, Dijkema et al. (1980)⁵¹ and Streif (2004)³ describe the existence of lagoons in the transgression cycle (peat - lagoonal clay - tidal flat sands with shells - lagoonal clay - peat) occurring in the Wadden Sea during the early Holocene, which supports the existence of lagoon. Today, this sedimentary environment has largely disappeared, but a few remnants can be found nearby the islands of Amrum and Trischen and near Sankt Peter-Ording⁵². Ehlers (1988)⁵² describes modern lagoons for the islands of Amrum and Trischen and near Sankt Peter-Ording.

Nevertheless, this facies interpretation needs to be confirmed by additional palaeontological data.

2.4.2.6. Sand flat deposits

The unimodal and wide shape of L* centred at high value of 17 (Figure 2-7B) reflects the pale colour containing grey sand and whitish shell fragments. The unimodal and narrow shape of both PC1 score and MS, centred at low values (Figure 2-7B), indicates the dominance of sand (mostly quartz) in the sediments. The high values of CT density, which are comparable to those of moraine, eolian/fluvial, and beach-foreshore facies, support this explanation. The high PC2 score (Figure 2-7B) implies the existence of carbonates (i.e., shells). These characteristics (few organic matter and rich in carbonates) agree with that as a tidal sand flat by Bulian et al. (2019)³⁰. However, this is not supported by distinctly higher CT-density data because the density of shell fragments (> 1700 HU) was excluded during CT-density data treatment.

2.4.2.7. Channel fill deposits

Since the subtidal channel receives sediments from both, sand and mud flats, additionally influenced by tidal movements, these deposits show bimodal characteristics (Figure 2-7C) supported by CT density and PC3 score. The unimodal and broad shape of the L* distribution (Figure 2-7C) indicates a darker colour compared to the sand flat, and lower colour contrast between mud and sand in the tidal area. Moreover, this low contrast may be induced by averaging data of the lenticular bedding due to the limits of the photograph (2D) while CT density has 3D data points. Slight increase of the PC1 score and decrease of the PC2 score compared to the sand flat facies imply a subtle signal change caused by the addition of organic matter and dilution of carbonates by the mud flat sediments. We only present one of the sections with this facies in Figure 2-3B. The remaining sections are shown in Figure S 2-2 and Figure S 2-3.

2.4.2.8. *Beach-foreshore deposits*

Brownish shell fragments mark the bimodal distribution of L^* (Figure 2-7D). The peaks represent dark grey sand and shell fragments (left to right), while the remaining results resemble the sand flat facies. However, if we look at the three PCs' relative positions (Figure 2-7D), we detect small differences. When taking PC3 as an anchor, the decrease of PC1 and the increase of PC2 scores imply that the deposition is less organic and more carbonaceous (marine), probably due to the lack of protection by the barrier island.

2.5. Limits and outlook

Although the combination of scanning results in density plots (Figure 2-6 and Figure 2-7) makes comparisons and interpretations convenient, this mixing of data may cause issues due to different applied techniques' scanning resolution. Scanning resolutions are defined by integration of two parameters, step size and window size. The step size determines the frequency of scanning points. If, for instance, the step size is larger than lamination thickness, the number of lamination is underestimated. In other words, the sampling rate is lower than the signal frequency. This phenomenon bridging the continuous and discrete signals is addressed by the Nyquist-Shannon theorem^{95,96}. On the other hand, the window size (e.g., the pixel size of images) determines each scanning point's scanning window. For example, a datapoint obtained with a larger scanning window than the thickness of lamination reflects only an averaged lamination value. With different settings of these two parameters, the resolution varies. Francus and Pirard (2005)⁹⁷ reveal that images taken with the same camera and from the same mineral thin-section but with different resolutions can generate a different phase. This issue occurs in lagoonal sediments of this study. Moreover, the averaging process from 2D and 3D to 1D data creates different levels of boundary blur, which has consequences for the characterisation of finely laminated deposits.

In fact, a sedimentary facies is rarely homogenous. There are variations of sedimentological and geochemical properties within each facies. For example, a moraine may contain sections of well-sorted sediments ¹. Peats can be composed of fen and peat bogs according to nutrient availability ⁹⁸. Soils have commonly different horizons resulting from soil-forming processes ⁹⁹. Due to the regional sampling scale of this study and a limited amount of core sections available for analyses, this study cannot include all variations in each facies. Grain size is one of the most important factors to determine sediment facies. This study includes the grain-size discussion based on lithological descriptions and PCA interpretations, but it is not supported by any conventional grain-size measurements. Consequently, facies are simplified to general categories instead of providing a detailed classification, which needs a spatially more diverse sampling database and more quantitative data, such as palaeontological and grain-size analyses. Nevertheless, our approach is the first practical step to analyse the data of different core scanning techniques to evaluate their potential for facies characterisation.

Machine learning has progressed considerably in various disciplines over the past two decades ²². Recently, there have been studies developing automatic facies classification on geological records. These deploy a vast variety of variables including seismic reflection data, petrophysical logging data, hyperspectral imaging or geochemical measurements to discriminate facies or classify sedimentary structures ^{25, 35-39}. These studies show that the application of machine learning in sedimentology is still at an early stage. Often, input data are limited and applications are restricted to a small study area or few cores since they are experimental studies. Our study contributes to broader facies coverage and data variety (element data, CT-density, L* and MS variations). The density plots combine these scanning data and illustrate the sediment characteristics of each facies visually. The differences among facies confirm the ability to reproduce the sedimentological observations in a digitised

perspective. Therefore, there is a high potential to develop automatic facies classification successfully in the near future.

The goal of machine learning is to train computers to use exemplary data for solving a given problem ²³. Based on sedimentological expertise, machine learning builds a facies discrimination model in the digital world. The scheme of developing an automatic facies classification using digitised information will be applied on a local scale with general facies classification first, and subsequently progresses to a larger scale and finer classification to cover a higher facies variability. With extended data coverage, inconsistencies and conflicts between results of different studies may appear. For instance, different facies recognitions are made based on different approaches. Nevertheless, we anticipate that future facies classification will be less time-consuming and allow sedimentologists to focus on other critical scientific questions.

2.6. Conclusions

This study demonstrates that eight ‘a priori’ recognised sediment facies (moraine, eolian/fluviol, soil, peat, lagoonal, sand flat, channel fill, and beach-foreshore) are (semi)quantitatively characterised and differentiated by their physical and chemical properties obtained by scanning techniques (MS, CT, μ -XRF, digital photography). The application of density plots offers a comprehensive way of presenting these properties and reveals distinct differences between facies. We document that conventional lithological descriptions can be reproduced by time- and cost-efficient (semi)quantitative and high-resolution digital measurements. However, some facies (lagoonal/salt marsh sediments and eolian/fluviol deposits) remain arguable, as scanning data show differences that were not previously recognized by conventional facies analysis. After all, digital scanning data is useful for further studies. Based on these results, an automatic facies classification model will be developed by applying machine learning techniques. In the end, a wealth of sedimentary information will be released from the shackles of expensive and observer-dependent data in the past.

2.7. Acknowledgements

The present study is part of the Wadden Sea Archive (WASA) project funded by the “Niedersächsisches Vorab” of the Volkswagen-Foundation within the funding initiative “Küsten und Meeresforschung in Niedersachsen” of the Ministry for Science and Culture of Lower Saxony, Germany (project VW ZN3197), coordinated by the Lower Saxony Institute for Historical Coastal Research (NIHK). We appreciate that the WASA project team jointly carried out the core recovery and macroscopic sediment description. ASL thanks the National Taiwan University (NTU) Research Center for Future Earth from the Featured Areas Research Center Program within the framework of the Higher Education Sprout Project funded by the Ministry of Education of Taiwan for travel support. JT received funds from the MARUM Cluster of Excellence ›The Ocean Floor – Earth’s Uncharted Interface‹ (Germany’s Excellence Strategy – EXC-2077 – 390741603 of the Deutsche Forschungsgemeinschaft DFG). We thank the Klinikum Bremen-Mitte and Prof. Dr. Arne-Jörn Lemke and Christian Timann for providing their facilities and supporting the performed computed tomography measurements.

2.8. Supplementary materials

The raw data can be found on <https://doi.pangaea.de/10.1594/PANGAEA.920071>.

Table S 2-1. General information about the studied sediment cores.

Core Section	Longitude	Latitude	Northing	Easting	Zone	Length (cm)	Facies
N14-2	7.13431	53.71048	5952926	376,864.37	32	100	channel fill, moraine, eolian/fluviol
N27-2	7.182342	53.6976533	5951418	379,997.65	32	95	channel fill
N33-2	7.217895	53.6938717	5950937	382,334.11	32	99	channel fill
N35-4	7.167479	53.6914642	5950754	378,998.73	32	108	sand flat
N41-4	7.232223	53.6931083	5950829	383,277.96	32	99	channel fill
N59-1	7.17137	53.75187	5957467	379,428.59	32	107	channel fill
N72-2	7.14738	53.6951296	5951196	377,682.42	32	109	eolian/fluviol
VVC16-3	7.130155	53.68284	5949859	376,509.28	32	117	soil, eolian/fluviol
VVC16-4	7.130155	53.68284	5949859	376,509.28	32	119	peat, lagoonal
VVC20-3	7.111546	53.7598445	5958457	375,508.10	32	25	beach- foreshore

Table S 2-2. Data points obtained by X-ray fluorescence core scanning (XRF), CT core scanning (CT), digital photography (L*) and magnetic susceptibility scanning (MS) for each sediment facies.

Sediment facies	XRF	CT	L*	MS
moraine	511	3878	24745	116
eolian/fluvial	814	5495	34397	166
soil	29	193	1234	6
peat	278	1992	12562	61
lagoonal	437	2972	19065	86
sand flat	506	3651	23070	110
channel fill	1992	13983	87735	415
beach-foreshore	116	857	5283	26

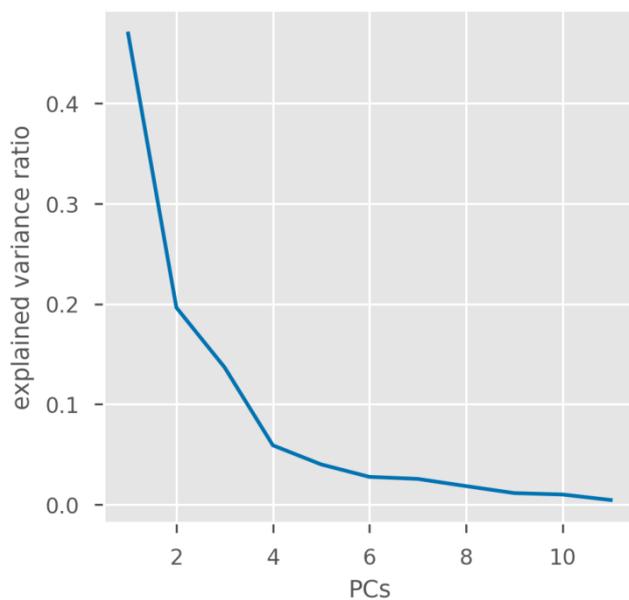


Figure S 2-1. Scree plot of the explained variance ratio for each principal component (PC).

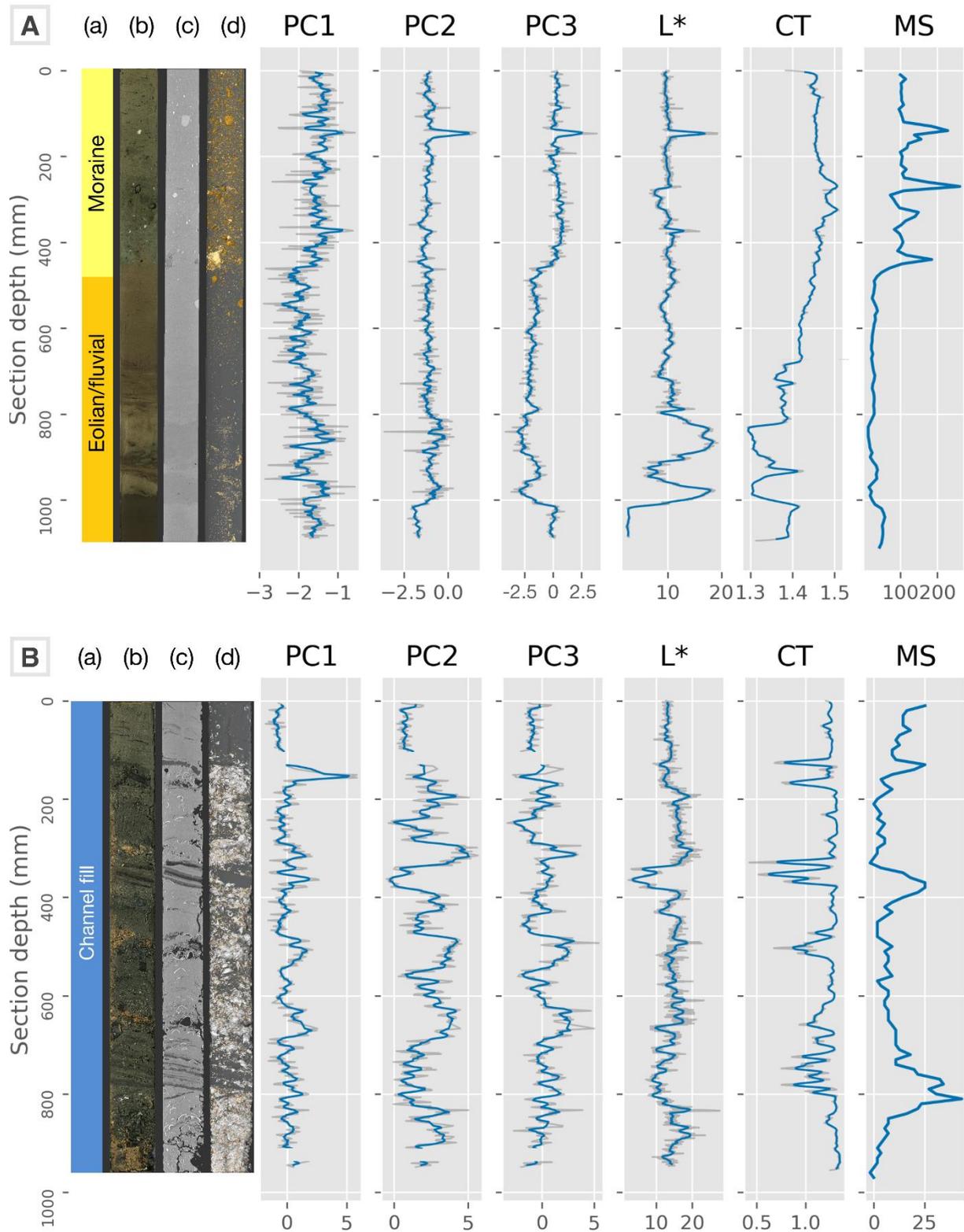


Figure S 2-2. Core N72-2 (A) and N27-2 (B): Facies labels (a), digital photograph (b), radiograph (c), and sediment constituents (golden: high-density sediments and lithic clasts, peach: low-density sediments, white: shell (fragments) separated using CT data (d)). Scores of PC1-PC3 (dimensionless) from elemental intensities, lightness (L^*), ranging from 0 (black) to 100 (white), CT density (CT in 10^3 HU), and magnetic susceptibility (MS in 10^{-6} SI). Blue lines represent an 8 mm moving average, while grey lines represent raw data. No moving average is applied to MS.

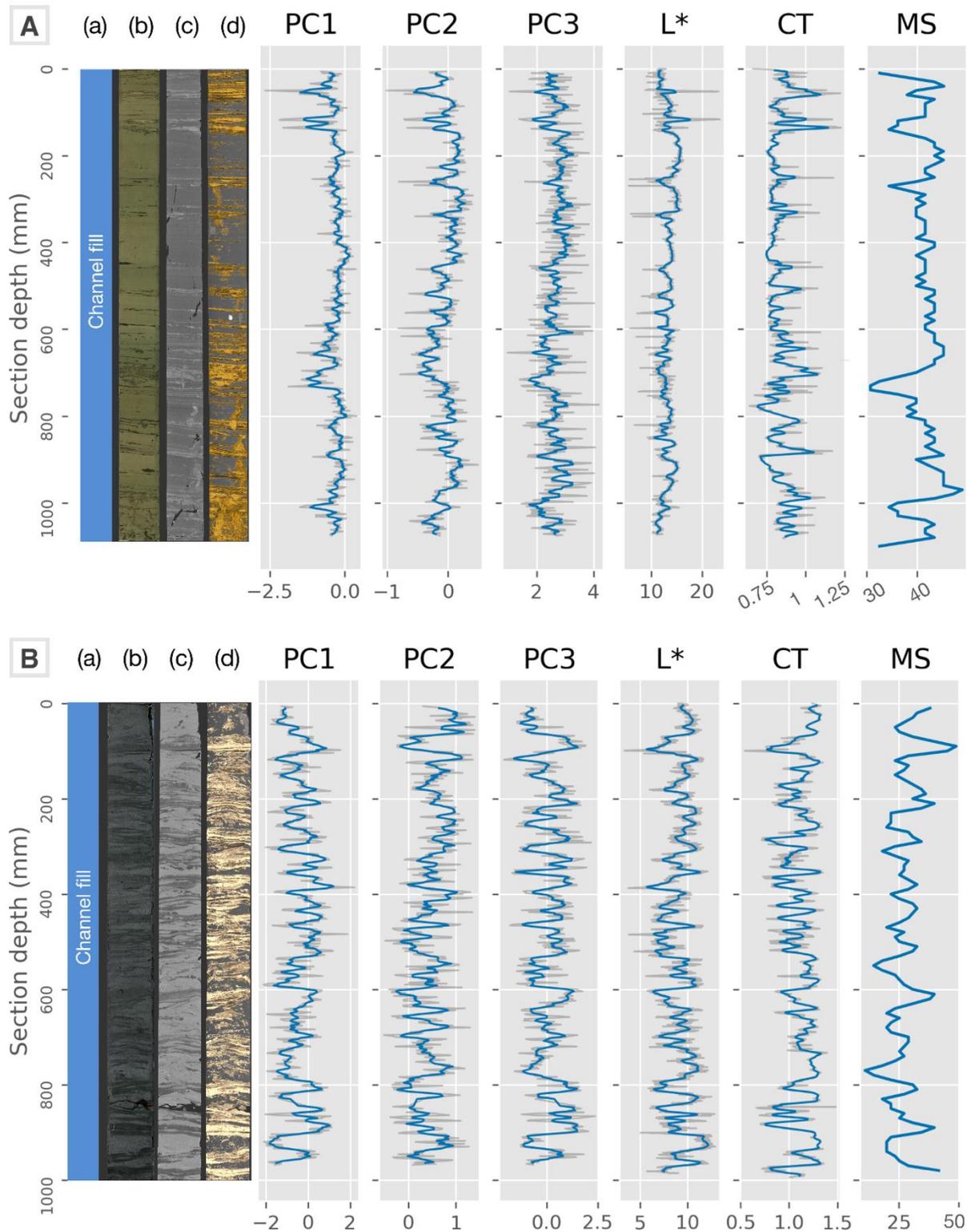


Figure S 2-3. Core N59-1 (A) and N41-4 (B): Facies labels (a), digital photograph (b), radiograph (c), and sediment constituents (golden: high-density sediments and lithic clasts, peach: low-density sediments, white: shell fragments) separated using CT data (d). The sediment constituents (d) reveals the intercalated layer structures and bioturbations. Scores of PC1-PC3 (dimensionless) from elemental intensities, lightness (L^*), ranging from 0 (black) to 100 (white), CT density (CT in 10^3 HU), and magnetic susceptibility (MS in 10^{-6} SI). Blue represent an 8 mm moving average, while grey lines represent raw data. No moving average is applied to MS.

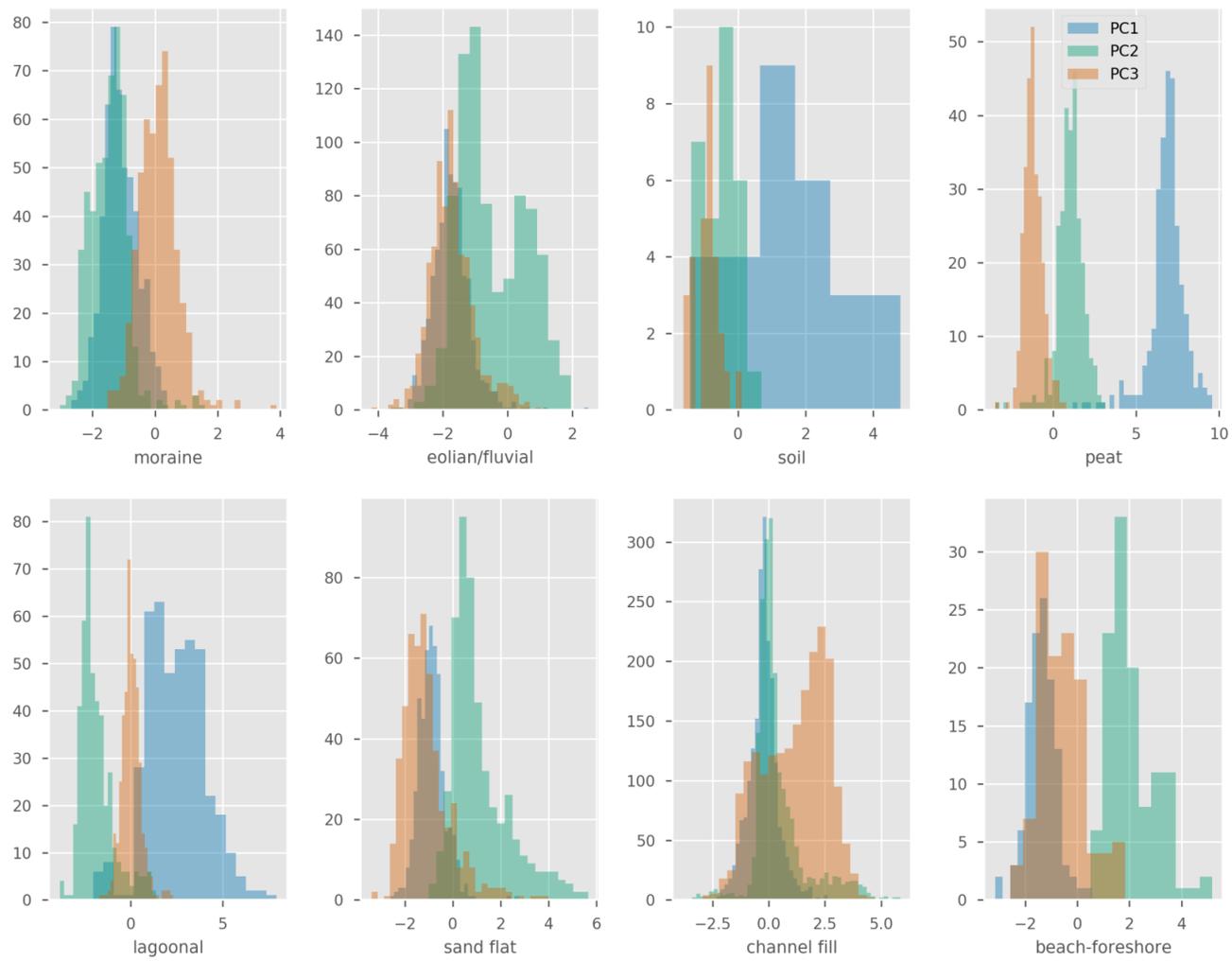


Figure S 2-4. Histograms of PC1 to PC3 scores for each facies. X-axis stands for the score value and y-axis shows the count.

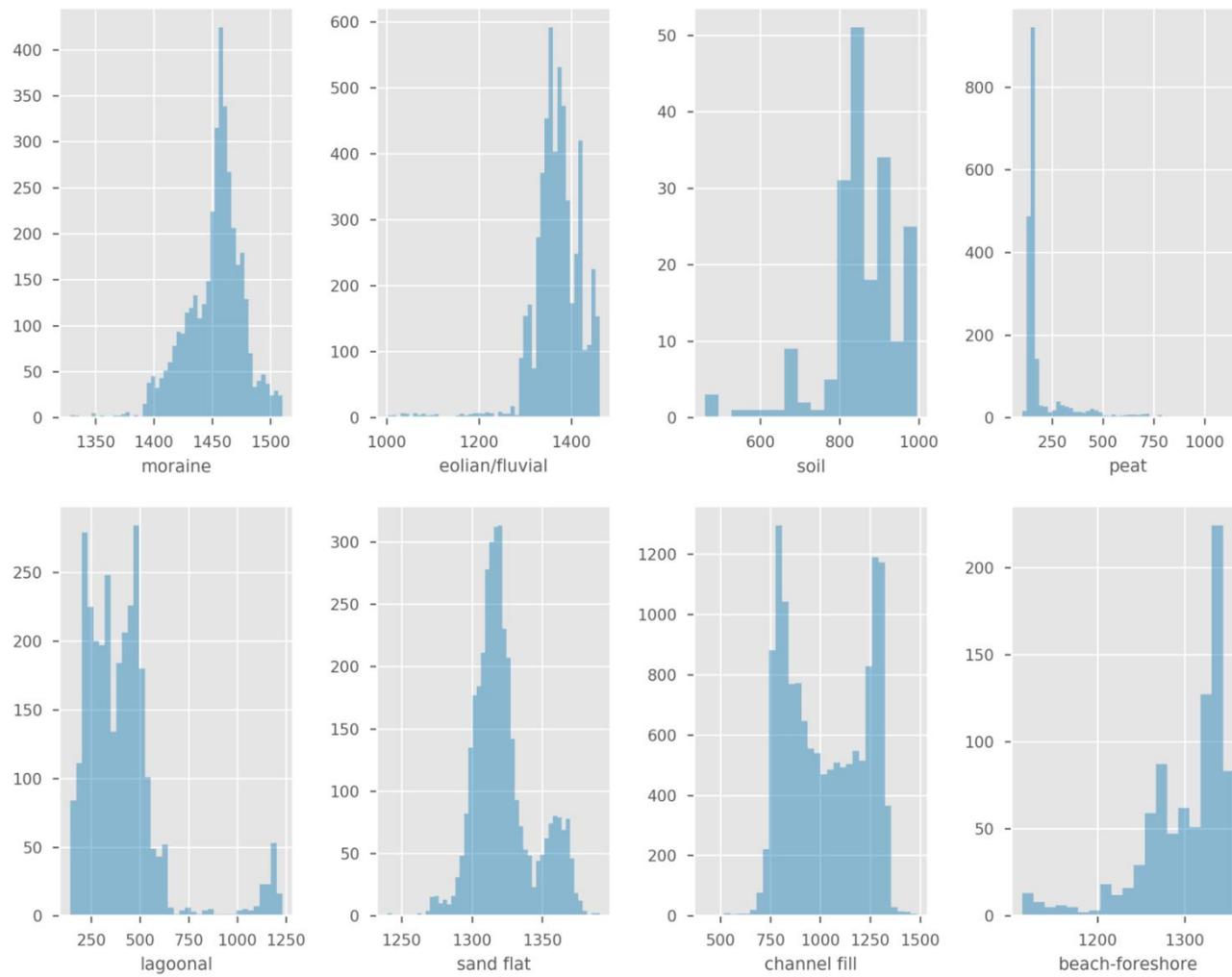


Figure S 2-5. Histogram of CT-density values for each facies. X-axis stands for the CT-density value (HU) and y-axis shows the count.

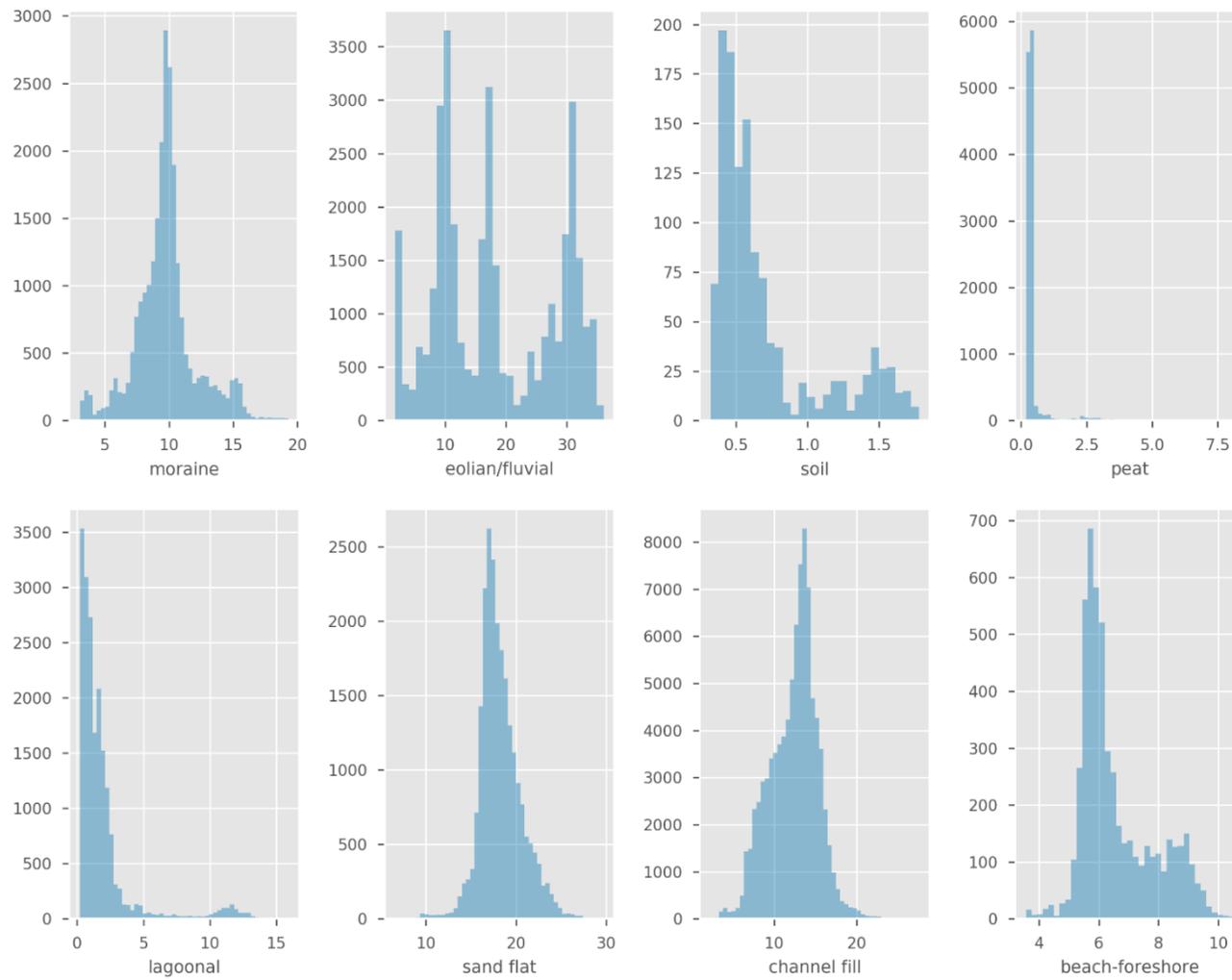


Figure S 2-6. Histogram of L^* values for each facies. X-axis stands for the L^* value and y-axis shows the count.

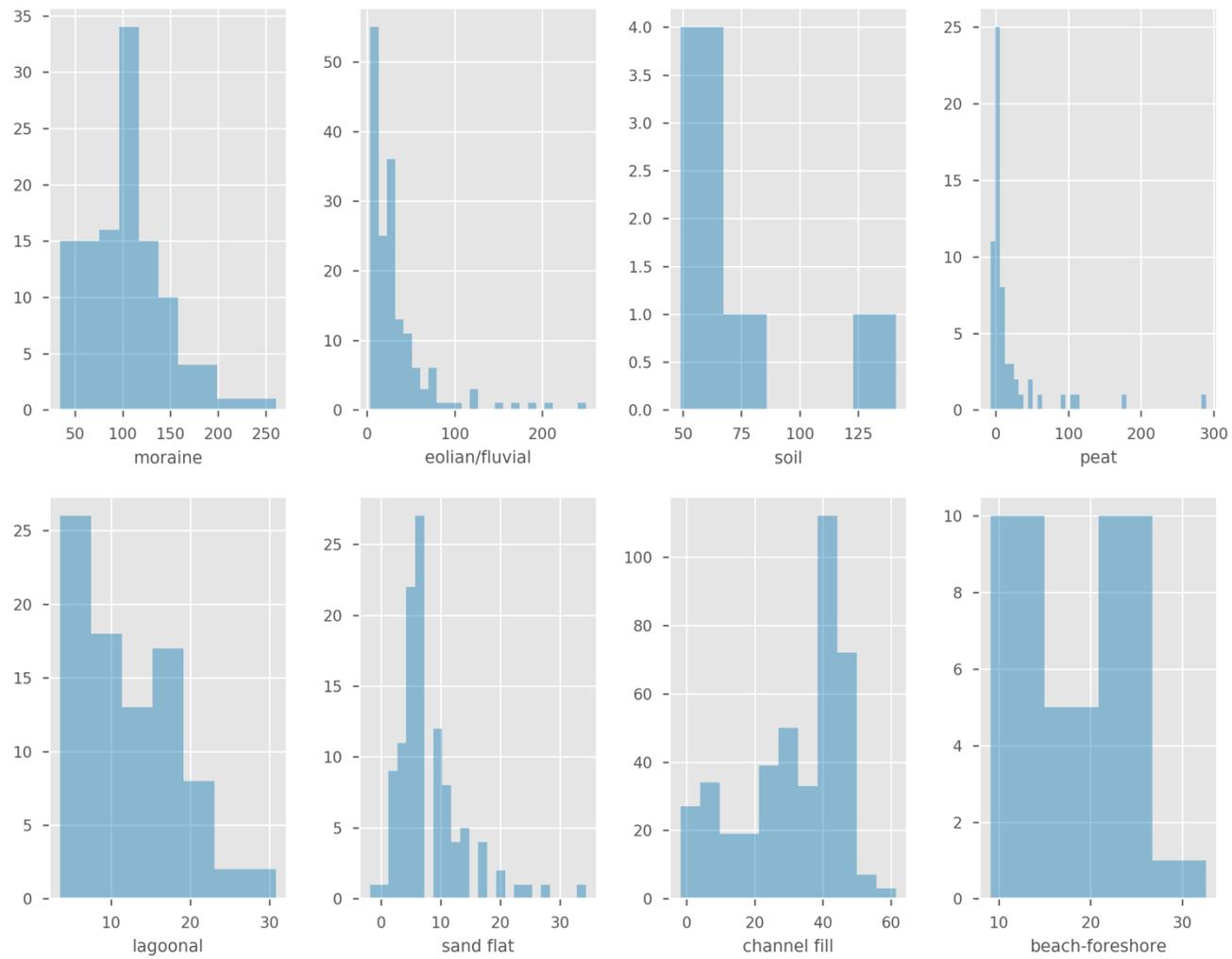


Figure S 2-7. Histogram of MS values for each facies. X-axis stands for the MS value (10^{-6} SI) and y-axis shows the count.

3. Study II: Automatic sediment facies classification

Progress in machine learning and sediment-core scanning creates a novel approach of automatic sediment-facies classification

An-Sheng Lee^{1, 2}, Dirk Enters³, Jyh-Jaan Steven Huang⁴, Sofia Ya Hsuan Liou², Bernd Zolitschka¹

¹University of Bremen, Institute of Geography, Bremen, Germany

²National Taiwan University, Department of Geosciences and Research Center for Future Earth, Taipei, Taiwan

³Lower Saxony Institute for Historical Coastal Research, Wilhelmshaven, Germany

⁴National Taiwan University, Institute of Oceanography, Taipei, Taiwan

In review

Communications Earth & Environmental

Author contributions

A. -S. L.: Conceptualization, Methodology, Investigation, Data curation, Writing – original draft.

D. E.: Conceptualization, Methodology, Writing – review and editing.

J. -J. S. H.: Writing – review and editing.

B. Z. and S. Y. H. L.: Supervision, Resources, Funding acquisition, Writing – review and editing

Abstract

The delineation of sediment facies provides essential background information for a broad range of investigations in the geosciences but is often constrained in quality or quantity. Two state-of-the-art evolutions in different disciplines (machine learning and sediment-core scanning) create innovative possibilities to cope with more comprehensive observations through computation, promoting a novel approach for automatic sediment-facies classification. This approach was developed using a regional-scale and high-resolution elemental dataset from sediment cores covering various sediment facies typical for the southern North Sea tidal flat, Germany. By accurately simulating the observational behavior of sedimentologists, an optimal machine-learning-built classification model (78% accuracy) involving a simple but powerful feature engineering is proposed. The model classifies the majority of sediment facies and highlights critical sections for further investigation. Research resources can thus be allocated more efficiently. Ultimately, it is expected to provide a generalizable blueprint to inspire more interdisciplinary research discovering findings at unprecedented depth.

3.1. Introduction

Sediment facies is defined as an assemblage of sediments that records specific characteristics of a depositional environment. It hence provides fundamental sedimentological information in space and time, which builds the necessary geological background for studying past surface processes on Earth ¹⁰⁰. For instance, the change of sediment facies through time assists scientists in reconstructing the evolution of paleoenvironments ^{3, 6, 101}. Moreover, the knowledge of sediment facies makes available a proper selection of geochemical proxies for a better interpretation of paleoclimatic variations ^{15, 42, 43, 102}. Exploring geological resources, such as for methane hydrates, oil and gas or offshore wind parks, also requires regional investigations of sediment facies ^{7, 8}. Basically, all research initiatives rely on the classification of sediment facies as a preliminary step.

Evolutions in computer science and computing power have brought machine learning (ML) techniques to various disciplines over the past two decades ²². Research has started to include ML to achieve in-depth scientific findings previously constrained by conventional methods and data resolutions ^{103, 104}. ML applications also have been introduced to the geoscience community, such as support vector machine, random forest and artificial neural networks ^{25, 35-38}. However, applications are still at an early stage. The sample size usually remains small ¹⁰⁵ and studies rarely apply a wide variety of classes (i.e., sediment facies or geochemical zones). Input data are mostly geophysical measurements (e.g., seismic profiles and borehole logging data) or laborious measurements (e.g., grain-size variation and quantitative element concentrations), which have resolution limits (tens of centimeters to meters in scale). Target classes are thus often homogeneous sediments (e.g., unique tephra layers) instead of complex units (e.g.,

laminated stream channel deposits). Furthermore, the ML models' decision commonly depends on individual data points rather than on comprehensive observations including adjacent data points, which compare much better to the way how sedimentologists work and investigate. All these factors limit its applicability; the same is true for the fact that developing codes are rarely open-source.

Evolution in core scanning techniques (e.g., X-ray fluorescence (XRF) core scanner, computed tomography and multi-sensor core logger) provides the possibility of acquiring near-continuous (μm -scale resolution), non-destructive and rapid measurements covering both geophysical and geochemical data from natural archives^{11, 12, 15}. A larger quantity of sediment investigation is thus integrated with higher spatial resolution and much more measuring details¹⁰⁶. Thus, data variety and size ascend to a new level, which helps resolve previously unachievable scientific questions (e.g., paleoclimatic variation and anthropogenic interaction with the natural environment)¹⁰⁷⁻¹⁰⁹. Moreover, this evolution facilitates the preservation of geoscientific knowledge since investigations are recorded digitally and promote the FAIR principles¹¹⁰ (especially reusability).

This approach aims at enhancing the automatic sediment-facies classification by overcoming most restrictions. This is inherited to the successes of two evolutions: ML and core scanning techniques. High-resolution elemental profiles obtained by XRF core scanning were acquired under the interdisciplinary WASA project¹¹¹ from the UNESCO World Heritage-listed Wadden Sea in Northern Germany. This region belongs to a dynamic geological setting, where the environment varied due to the rapidly rising sea level during the Holocene from glacial (terrestrial) to shallow marine^{3, 92, 112}. Thus, the studied sediments provide comprehensive geochemical signals and broad coverage (11

sediment facies, 53 core sections, 19k data points, for details: Figure 3-1 and Methods), spanning the Late Pleistocene throughout the Holocene⁹². Instead of effortlessly applying ML algorithms, we tested different feature engineering methods (data transformations, principal component analysis - PCA) to simulate the sedimentologists' observational behavior by extending the model's single-point analysis to a multi-point analysis. Simple (logistic regression - LR) and complex (kernel support vector machine - SVC, random forest - RF) ML algorithms were included to compare their feature engineering-integrated performance. Careful evaluations involving cross-validation (CV) and tailor-designed measures were carried out during these steps to guarantee the generalization of the models.

The results indicate a noticeable benefit of feature engineering that empowers the use of simple ML algorithms, giving enhanced performance (78% accuracy) and applicability to a standard computer. In addition, the optimal model provides an outcome with a confidence level, which highlights critical parts of the sediment records to allocate more resources for analyzing and saving resources from the remaining sediments. The approach is expected as a blueprint to inspire further studies in geosciences regardless of data type, region and objective. More importantly, this is bringing interdisciplinary research to a new era based on state-of-the-art technological innovations.

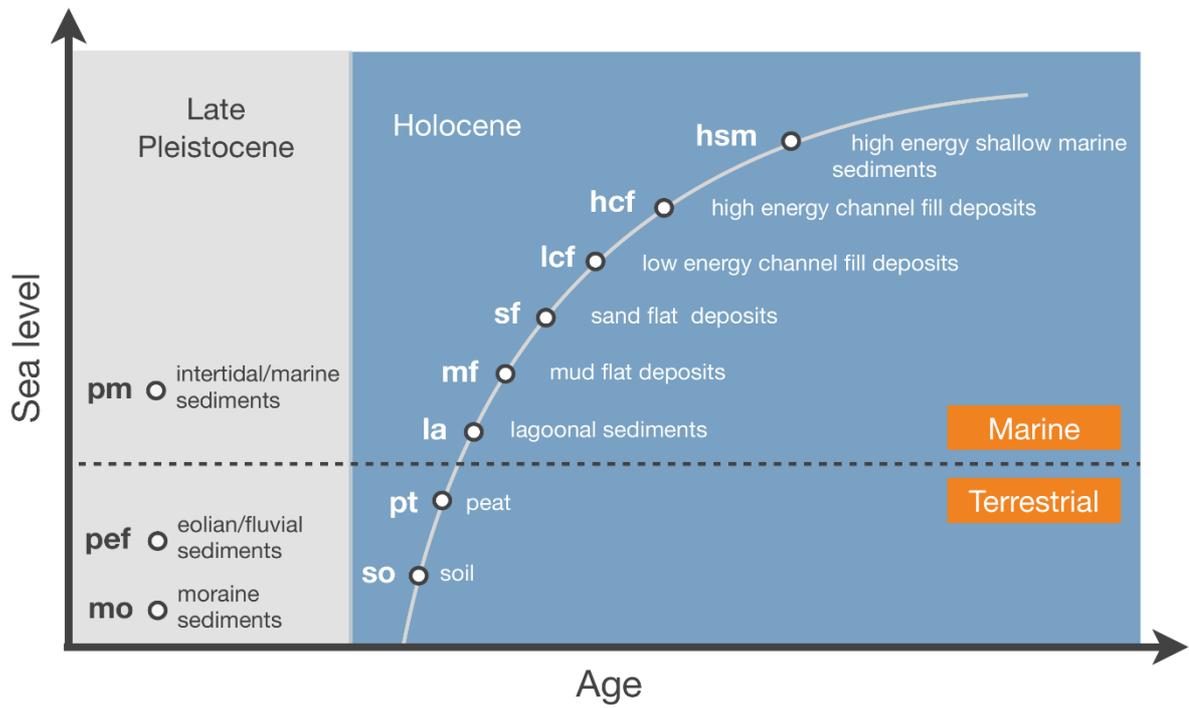


Figure 3-1. Information (abbreviation, age and sedimentary environment) for facies occurring in the study area, a tidal flat in the northern German Wadden Sea (adopted from Capperucci et al. (2022)⁸⁴). The positions of the facies in relation to sea level and age are only for illustration. For details see: Table S 3-1.

3.2. Results

3.2.1. Contribution of feature engineering

A series of models were built to classify sediments using the high-resolution elemental profiles automatically. The performances from each combination of feature engineering and ML algorithms were evaluated by the mean accuracy of 5-fold cross-validation (CV) scores. The output and visualization of the grid search are presented in Supplementary Figure S3-5. The highest CV scores for each combination are illustrated in Figure 3-2a. The CV scores increase after data representations, and the rolling representation results in best scores. Only 5 out of 9 comparisons between using or not using PCA transformation indicate “including PCA” as the better feature engineering. Thus, there is no clear advantage of using a PCA transformation. After data representation, the complex ML algorithms (SVC and RF), which have the enhanced learning power of non-linearity, provide no noticeable advantage over the simple algorithm (LR).

Furthermore, the best models using each kind of data (Figure 3-2a, raw data: SVC with PCA, rolling data: LR without PCA, 2D data: LR with PCA) were evaluated in the test set. Figure 3-2b shows that the model using rolling data provides the highest accuracy (78%) and fewest boundaries (98). Its boundary amount is about 20 times of that given by sedimentologists (5), which is still significantly lower compared to the model using raw data output (591). The model using 2D data does not perform as well as the one with rolling data but is still better than raw data. When considering the large size of 2D data (9 times of rolling data), creating a burden for computing, the preference for the rolling data approach is confirmed.

The optimal model built on the best ML combination (LR learned from rolling data without PCA) indicates its increasing performance on laminated facies. When

comparing its accuracies of channel fill deposits (especially hcf, Figure 3-2c), to those of the model using raw data (Figure 3-2d), the misclassification to sand and mud flats is remarkably reduced. This points out that the rolling representation well describes the laminated characteristic of channel fill deposits, comprising two kinds of homogeneous sublayers (sand and mud). The optimal model does not classify those sediments separately but together as one facies to recognize a specific deposition environment. The simple algorithm is thus allowed to surpass more complex algorithms to the benefit of feature engineering. This finding offers an efficient way of applying automatic sediment-facies classification by demanding less computing power due to less model complexity.

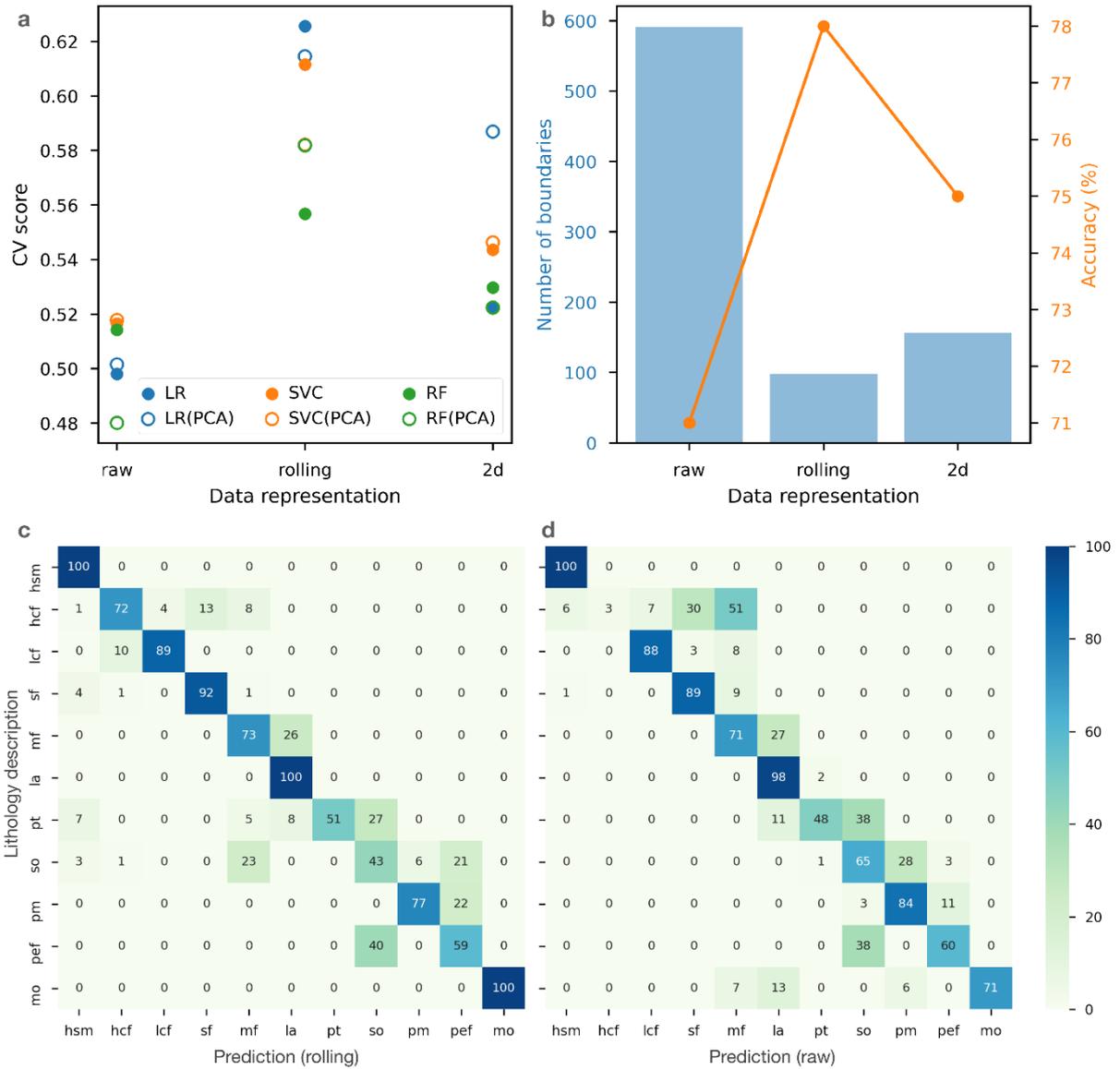


Figure 3-2. (a) The best cross-validation (CV) score of each combination of feature engineering (data representations and principal component analysis - PCA) and ML algorithms (logistic regression - LR), kernel support vector machine - SVC, random forest - RF - classifiers). The CV score is relevant to the mean accuracy during CV iterations. (b) The fragmentation, quantified by the number of boundaries, and accuracy of the optimal model built from each data representation in the test set. (c-d) The modified confusion matrices describe the performance of the optimal models built from (c) rolling data and (d) raw data when applied to the test set. The y-axis represents the facies from lithological description by sedimentologists. The x-axis stands for the model-classified facies. The numbers represent the percentages of data in each row, i.e., recall in statistical terminology. For instance, at the top left of (d), only 3% of data points recognized as hcf by sedimentologists are correctly classified by the model while 51% of these data points are misclassified as mf by the model. For abbreviations refer to Figure 3-1.

3.2.2. Misclassification of the model

Besides the improvements that benefited from feature engineering, the optimal model has limitations causing misclassifications. Error analysis summarizes four main error-causing categories (Figure 3-3a). The first category with 63.6% occurrence frequency describes the situation that the predictions could be correct, but this facies change is omitted to fit a general picture of environmental interpretation by sedimentologists. This is because the capture of composite characteristics is limited by the initially defined (fixed) window size during the rolling representation. For those laminations or sediment sections having a thickness >32 mm (window size: 17 data points), the optimal model cannot identify them as composite facies. These misclassifications are often considered as minor sedimentary structures. For instance, minor low energy channel fill deposits are misclassified to high energy channel-fill deposits in core section N31-1 (Figure 3-3b), which might reveal a small-scale channel deepening. In fact, sedimentologists have a flexible observation window to identify facies, but our ML-based optimal model can only start from a local measurement and joins with nearby information. Therefore, these misclassifications should be considered as “the model looks too detailed” rather than classification problems.

Another error category is the transition boundary problem, having 41.2% occurrence frequency (Figure 3-3a). Figure 3-2c shows that the accuracies in peat, soil, Pleistocene eolian/fluvial deposits and Pleistocene marine sediments are relatively low compared to the other facies. In most cases, the boundaries between these facies are gradual and/or indistinct. It is difficult for sedimentologists to draw a clear separation between them via macroscopic observation and elemental variation. For example, soils developed from Pleistocene sediments, which makes them similar in their elemental

composition and causes confusion to the model. This cause of misclassification can be found in the boundaries of pt-so-pef (lower part of N71-4, Figure 3-3b).

The third category having 21.4% occurrence frequency (Figure 3-3a) is related to the limitation of macroscopic observation by sedimentologists. Sedimentologists describe these misclassified sediments with doubts since they are visually similar to several sediment facies. For instance, in N18-2 (Figure 3-3b) around 0.6 m depth, structureless sediments are described as eolian/fluvial deposits like its neighbors by sedimentologists. However, according to their elemental profiles, the optimal model classified them differently from the neighbors as Pleistocene marine sediments. This is reasonable because the Pleistocene sand flat sediments might be visually identical to the structureless eolian/fluvial deposits unless checking their diatom and geochemical data.

Subjective and opposing judgments also happen between sedimentologists. Our model provides a consistent judgment through its fixed and relatively small observation window. Only a few cases like wood fragments are misclassified as high energy shallow marine sediment, are out of the model's ability. This material (3.7% occurrence frequency, Figure 3-3a) is not in the facies list of training data for model building. The detailed error-analysis result and comparison results are listed in Table S 3-2, Figure S 3-6 and Figure S 3-7.

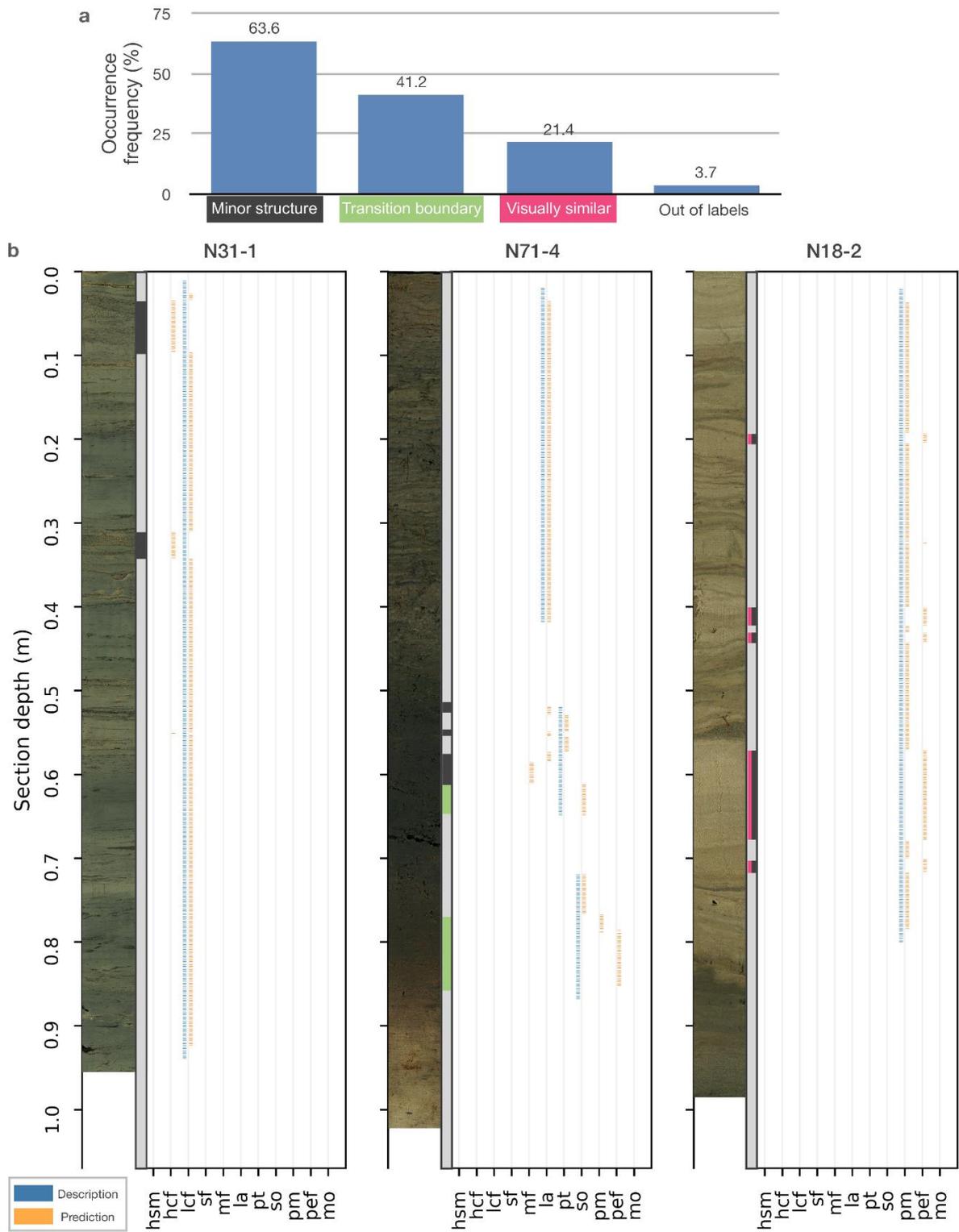


Figure 3-3. (a) Error categories and their frequency in misclassifications of the test set. The sum of the frequency is not 100% because categories may co-exist. (b) Test set comparison between sedimentologists' descriptions (blue) and predictions (orange) of the optimal model built using LR from the rolling data without PCA transformation. The core images and error category labels (marked in the color code of (a)) are attached. For abbreviations refer to Figure 3-1.

3.2.3. Highlighting critical segments

Our optimal model estimates the confidence level of each classification. Figure 3-4a demonstrates an example of the probability distribution along the core depth of one selected core section. If the model has low confidence in its classification, the probability values are dispersed along the facies. In contrast, the high probability in a specific facies stands for the model's confident decision. To more conveniently recognize the confidence level of our model, the maximum probability among facies for each data point is extracted (Figure 3-4b). The higher the maximum probability value, the more confident the decision is.

The relation between the model's confidence level and its error rate is illustrated in Figure 3-4c. The misclassifications (labeled as incorrect) mainly have a low maximum probability, suggesting that the more uncertain our model is, the more likely it is wrong. Empirically, the model's classification has a higher chance of being wrong when the maximum probability value is <0.3 . Therefore, sedimentologists can focus on sediment sections with a low maximum probability to carry out further examinations, such as microscopic observations or quantitative geochemical analyses, while adopting the automatic classification done by the model if the maximum probability is >0.3 . This provides a valuable outcome by highlighting the critical part where the expertise and experience of sedimentologists are needed for correct facies classification. When applying the model to the whole dataset of WASA (all 92 cores, 383 sections, 159k data points), only 1% of the sediments are marked as critical requiring further analysis.

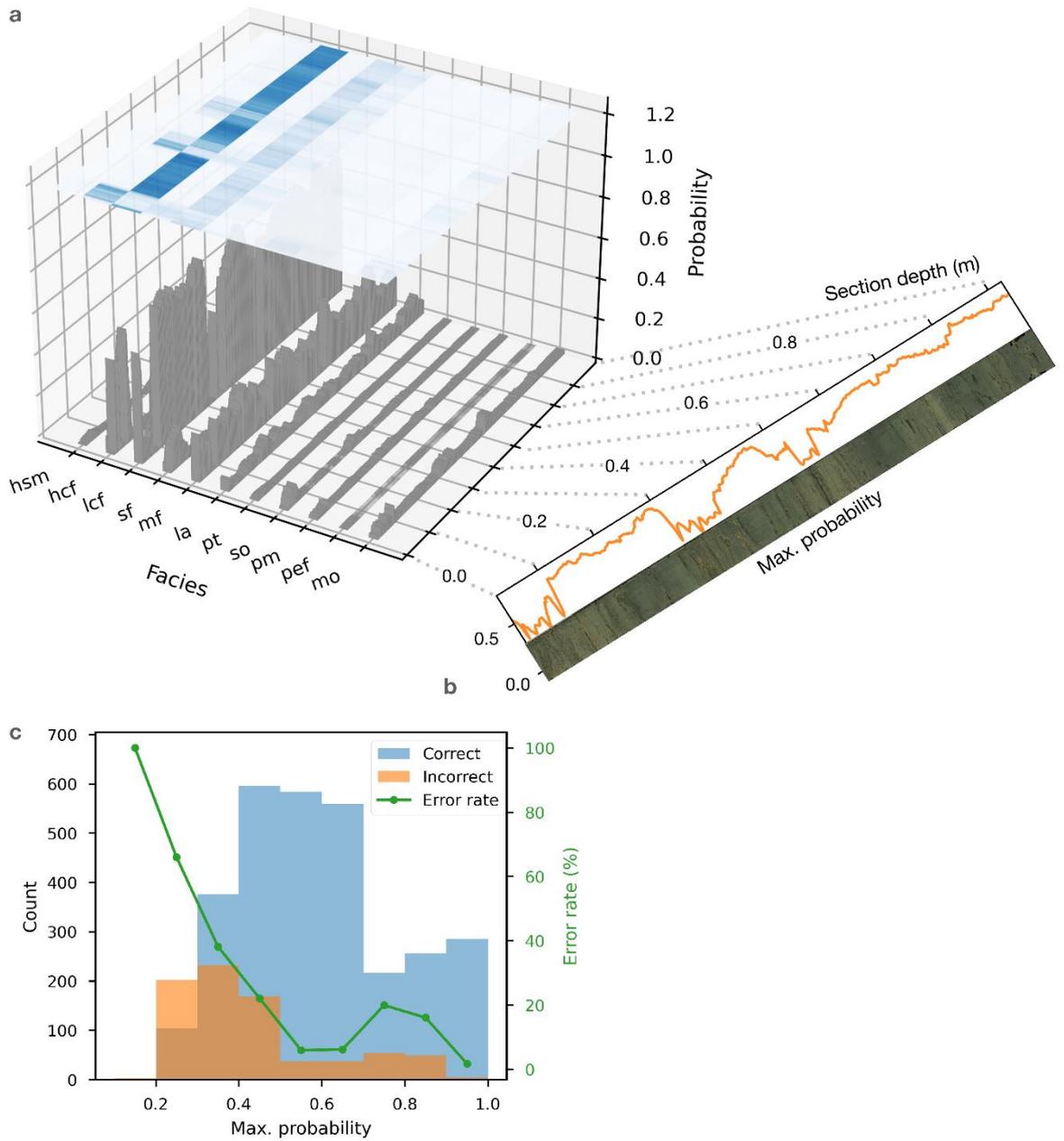


Figure 3-4. (a) Probability along core section N31-1. The probability of each facies in each data point is visualized in bars and projected to a heatmap above (the deeper the blue, the higher the probability). (b) The maximum probability for each data point is plotted along with the core image. For abbreviations refer to Figure 3-1. (c) Bi-plots of (1) distribution of the maximum probability in the groups of correct and incorrect classifications and (2) error rate for each interval of the maximum probability in the test set.

3.3. Discussion

Our study provides not only a model to automatically classify sediments from the Wadden Sea in Germany, but can also be considered as a methodological blueprint for projects, regardless of the type of data and the region investigated. Compared to other studies developing automatic classification of geological records^{25, 36, 37}, our approach has elevated the application to a new level, which is neither restricted by the dimension of input data nor its spatial coverage and facies variability. A simple but powerful recipe of feature engineering (centred log-ratio transformation and rolling representation) is proposed to fit the specific needs of sedimentological applications after testing several combinations of feature engineering methods. The best model does not need complex ML algorithms (SVC and RF), thus reducing computation time and model size.

A possible reason explaining the usefulness of rolling representation is its combination of measurements (i.e., geochemical properties) and heterogeneities from the sample and adjacent sediments. This data representation moves from a single-point observation to a multi-point (window-sized) observation, which simulates human observation by considering the context of sedimentary features to define the sediment facies. The rolling mean draws an overall elemental fingerprint of the sediments, while the rolling standard deviation describes heterogeneity within this fingerprint.

The comprehensive evaluation of the model is estimated by using robust cross-validation in the training set as well as boundary amount and balanced accuracy in the test set. Based on a performance of 78% accuracy, most misclassifications – often described as unavoidable bias¹¹³ – are discussible and acceptable. More importantly, the quantified confidence level helps sedimentologists invest more effort in the critical parts of the sedimentological record by saving resources from rather simple but often vast

sediment sections. This progress may redistribute resources more efficiently in future studies.

As mentioned for the model's limitations, our approach has not yet fully simulated the sedimentologist's behavior. With a flexible observational window and a bigger picture of environmental interpretations in mind, the complex or customized architecture of neural network algorithms, like long short-term memory, may cause improvements. In addition, elemental profiles may not be sufficient. The data dimension could be enlarged, e.g., by including image data or developing a higher data coverage. Also, more sediment facies can be considered, which should boost the success of automation.

3.4. Methods

3.4.1. Data acquisition

This study was developed under the framework of the interdisciplinary Wadden Sea Archive (WASA) project ¹¹¹. Ninety-two sediment cores (length: up to 6 m, diameter: 8 and 10 cm) were recovered from tidal flats, channels, and offshore around the island of Norderney. A team of sedimentologists carried out lithological descriptions and sediment-facies interpretation through macroscopic observation ⁹². High-resolution elemental profiles were acquired from a COX Itrax-XRF core scanner at the GEOPOLAR lab, University of Bremen, scanned at a fixed setting and subsequently processed by the Q-spec software (version 2015; COX Analytical Systems). Twelve elements (Si, S, Cl, K, Ca, Ti, Fe, Br, Rb, Sr, Zr, Ba) were chosen based on signal reliability. All data and information were adopted from a previous study ¹¹⁴ that compiled the geochemical and geophysical measurements of the sediment cores. Fifty-three representative sediment core sections (length: <1.2 m, locations: Figure S 3-1a, with 19,823 data points) covering 11 sediment facies (Figure 3-1) were selected for this approach in order to constrain computing time. Data points representing cracks and rough or uneven sediment surfaces were excluded. The facies labels extracted from lithological descriptions and elemental profiles were aligned according to their depth.

3.4.2. Feature engineering

The developing scheme of this study is demonstrated in Figure S 3-2a. The dataset was randomly split into training and test sets for a robust evaluation. The test set contains one section for each facies while the remaining sections were used for training. A series of feature engineering methods (i.e., data transformations in ML terminology) were deployed to enhance the performance of the models built by supervised ML algorithms.

Then, combinations of these methods were tested to find the most useful one. The steps are the following:

Elemental data were normalized by the geometric mean in each data point to eliminate the variance caused by the machine and the sediment itself, such as XRF tube aging, water content and grain size, to achieve a better prediction. In addition, the logarithm was applied to free the normalized data from asymmetry and closed-sum effects^{115, 116}. Together, this data treatment is called centred log-ratio transformation, which is commonly applied to compositional data⁶⁸.

To capture the composite characteristic of sediment facies, the neighboring information of each data point needs to be included to the data point. We propose two data representations for comparison. The first is rolling representation, where the dataset is represented by a centered moving mean and its standard deviation (Figure S 3-2b) for each element. The moving window size was set to 17 data points since the step-size resolution of the Itrax XRF core scanner is 2 mm and the thickness of beddings is predominantly <1 cm for studied sediments¹¹⁴. The second is called 2D representation. It collects adjacent data points as a chunk of data (17 data points * 12 elements) and raveled them to an array of new dimension (Figure S 3-2c). This approach is common in image analysis, which transforms a 2D pixel data matrix to a 1D array. These representations were implemented for each core section individually to prevent the model building from data snooping¹¹⁷.

In the next step, the represented dataset was standardized to zero mean and unit standard deviation, which is essential for some ML algorithms. Principal component analysis (PCA) with correlation matrix and whitened settings was included in the workflow to discuss its need in feature engineering. Yet, the standardization was not

applied to the dataset for the combination of random forest classifier (RF) without PCA because RF is not sensitive to the variable's scale difference.

3.4.3. Model building and evaluation

After feature engineering, three algorithms were applied to learn the dataset and to build models to classify sediments into facies by analyzing elemental profiles automatically. Logistic regression classifier (LR) using 'lbfgs' as solver ¹¹⁸ and L2 regularization was selected from linear algorithms. Kernel support vector machine (SVC) and RF classifiers ^{119, 120} were selected from sophisticated non-linear algorithms. The kernel technique of SVC allows exploring data relations in the infinite space ¹²⁰, which includes element ratios. The algorithms were implemented with the balanced class (i.e., facies) weights ¹²¹ to deal with our imbalanced facies distribution (Figure S 3-1b).

Prior to the final evaluation, an exhaustive search for parameters using 5-fold cross-validation (CV) was utilized to (1) assess the usefulness of data representation, (2) evaluate the need of using PCA and (3) fine-tune parameters (LR: C, SVC: C and γ , RF: n_estimators and max_depth) for building models. In brief, those parameters control the regularization strength of an algorithm to achieve a better generalization ²¹.

During each iteration of CV, the training set was split into two sets (2nd training and validation sets) following the splitting strategy mentioned above. The 2nd training set was learned by a pipeline composed of feature engineering and algorithm with specific parameters to build a model. The model's performance was evaluated by its accuracy for the validation set. Instead of using simple accuracy, which calculates the overall percentage of the correct predictions, we used "balanced accuracy", a macro-average of recall scores per class or, equivalently, the mean of accuracies from 11 facies ^{122, 123} for evaluation. This helps with the search for optimal parameters that build models predicting

all facies fairly well rather than predicting the dominant facies only. A mean validation accuracy (CV score) was calculated after iterations, which is more robust in facing data heterogeneity. In total, 18 CV scores were collected from the best CV scores of each combination (3 kinds of data, with or without PCA, 3 algorithms).

Once the optimal parameters were decided based on CV results, these algorithms learned the whole training set to build optimal models. The test set was then used for evaluating the models' performances. The performance was described by balanced accuracy, demonstrated in a modified confusion matrix, and the number of boundaries, describing the fragmentation problem of predicted facies.

Error analysis was also carried out on the test set to investigate the underlying causes of misclassifications and consequently prioritize possible improvements¹¹³. Misclassifications were thoroughly checked via both elemental profiles and images to summarize error categories. Each misclassification was given one or multiple error categories. In the end, the occurrence amount of each category was divided by the total amount of misclassification as frequency, i.e., the category's contribution.

All computations and visualizations were conducted using the SciPy ecosystem in Python^{73, 79, 80, 83, 124, 125}. The function of providing model probability was adopted from the well-developed Scikit-learn package⁷³.

3.5. Acknowledgements

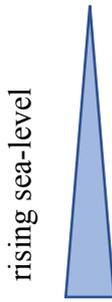
The present study is jointly funded by (1) the Wadden Sea Archive (WASA) project funded by the ‘Niedersächsisches Vorab’ of the VolkswagenStiftung within the funding initiative ‘Küsten und Meeresforschung in Niedersachsen’ of the Ministry of Science and Culture of Lower Saxony, Germany (project VW ZN3197), and (2) the Ministry of Science and Technology of Taiwan (project number: MOST 110-2116-M-002-023). We appreciate the WASA project team that jointly carried out the core recovery and macroscopic sediment description. A. -S. L. thanks the National Taiwan University (NTU) Research Center for Future Earth from the Featured Areas Research Center Program within the Higher Education Sprout Project framework funded by the Ministry of Education of Taiwan. We acknowledge Ben Marzeion and Timo Rothenpieler (Klimageographie, University of Bremen) for generously providing high-performance computing power.

3.6. Supplementary materials

Data Will be available on [Pangaea](https://pangaea.de). All computer codes and instructions are open-access on https://github.com/dispink/WASA_faciesXRF.

3.6.1. Material and methods

Table S3-1. Description of sedimentary facies occurring in the sediments of the northern German Wadden Sea (adopted from Capperucci et al. (2022) ⁸⁴).

Epoch	Facies	Abbreviation	Sea level	Remarks
Holocene	high energy shallow marine	hsm	marine	subtidal, coarse-grained sediments, e.g. shoreface, foreshore and channel lag
	high energy channel fill	hcf		subtidal, laminated coarse-grained sediments
	low energy channel fill	lcf		subtidal, laminated fine-grained sediments
	sand flat	sf		intertidal environment
	mud flat	mf		intertidal environment
	lagoon	la		brackish environment/salt marsh
	peat	pt		terrestrial
	soil	so	terrestrial	developed on Pleistocene material
Late Pleistocene	intertidal/marine	pm	marine	marine sediments
	eolian/fluviol	pef	terrestrial	structureless / unclear
	moraine	mo	terrestrial	glacial sediments

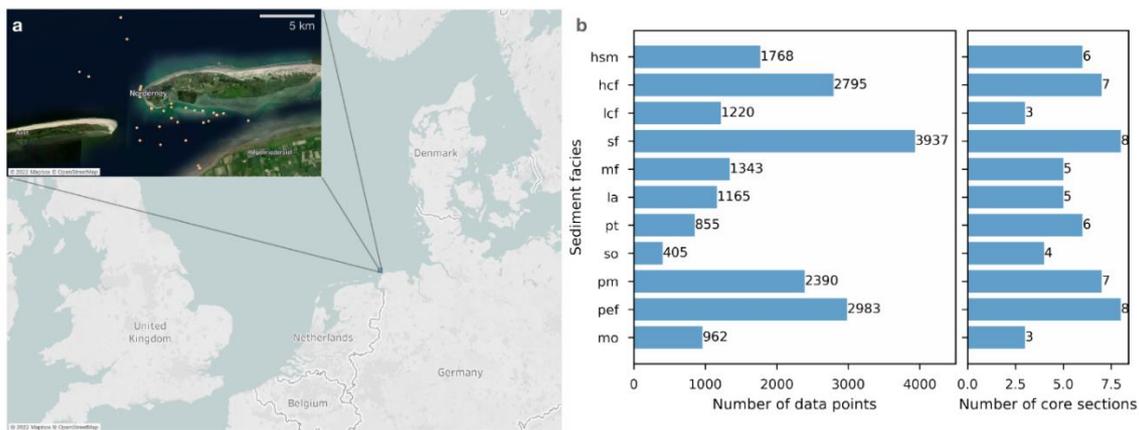


Figure S 3-1. (a) Location of coring sites in northern Germany with selected coring locations (orange dots in inserted map). (b) Number of data points and core sections for each facies. There are 19,823 data points and 53 sections in the selected dataset. Some sections have multiple facies. For abbreviations refer to Table S 3-1.

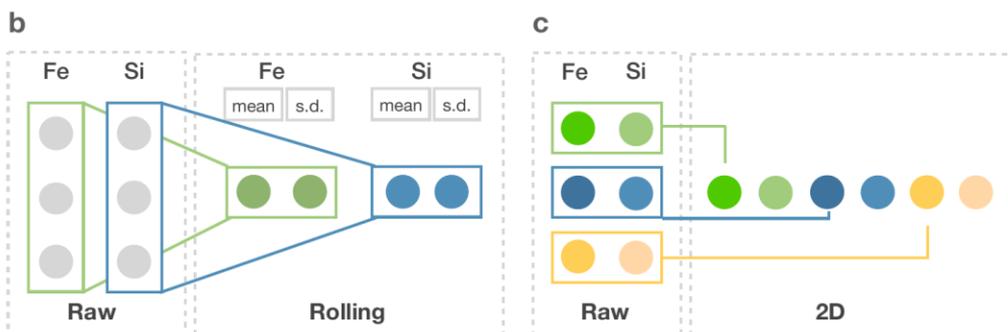
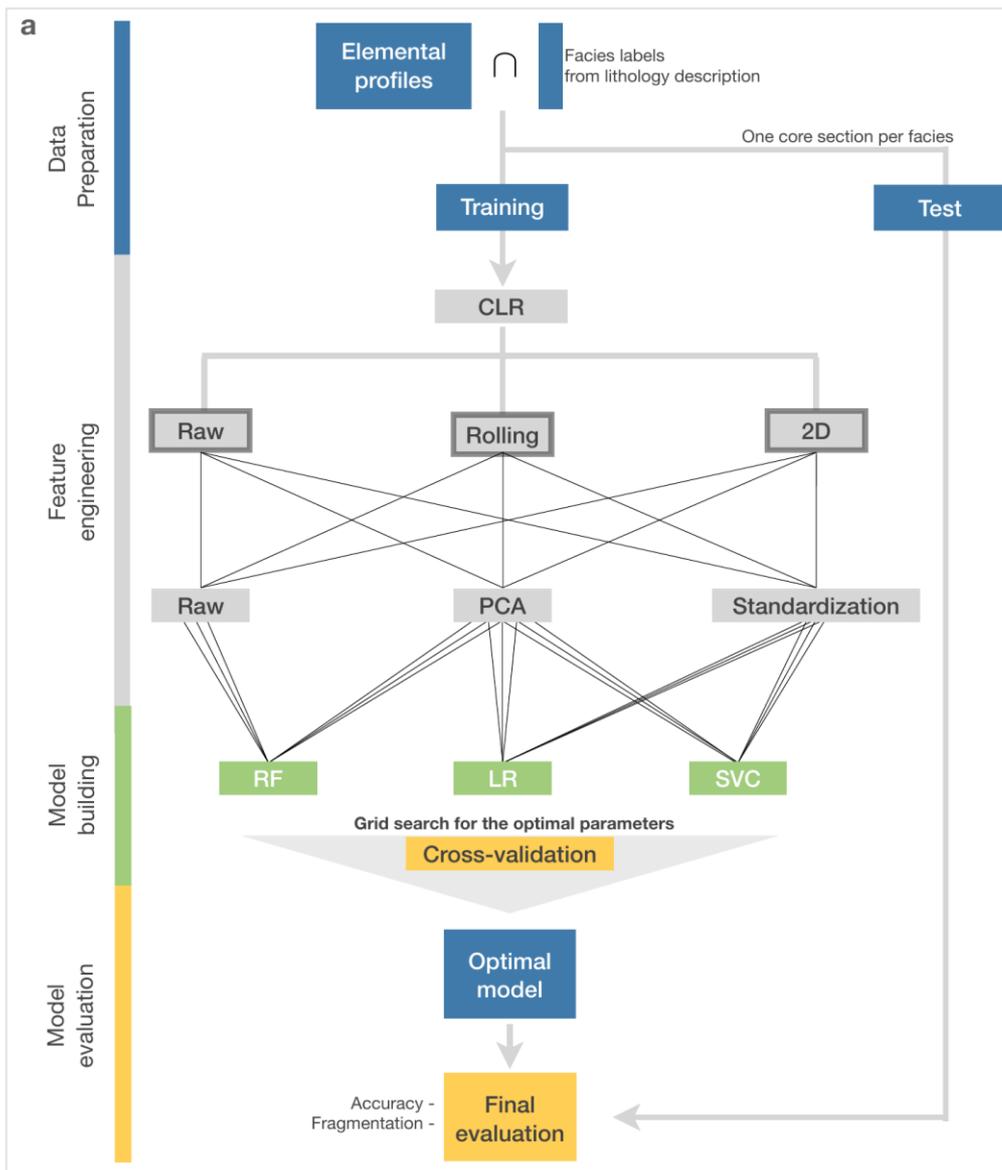


Figure S 3-2. (a) Developing scheme of this study. Illustrations of data representation for (b) rolling and (c) 2D approaches. The exemplary moving window contains raw data points of two element intensities (Fe and Si) in three depths. In (b), the raw data points are represented by the rolling mean and standard deviation (s. d.) of adjacent data points. In (c) the raw data points are represented by directly including the adjacent data points like a 2D image and raveled.

3.6.2. Grid searching results

The grid searching for optimal parameters of 18 combinations (3 kinds of data, with or without PCA, 3 supervised ML algorithms) are visualized in Figure S 1-3. The table of the compiled grid search results is stored as `r_grid_all_20221020.csv`. The optimal parameters are located either at the peak or plateau of the CV scores.

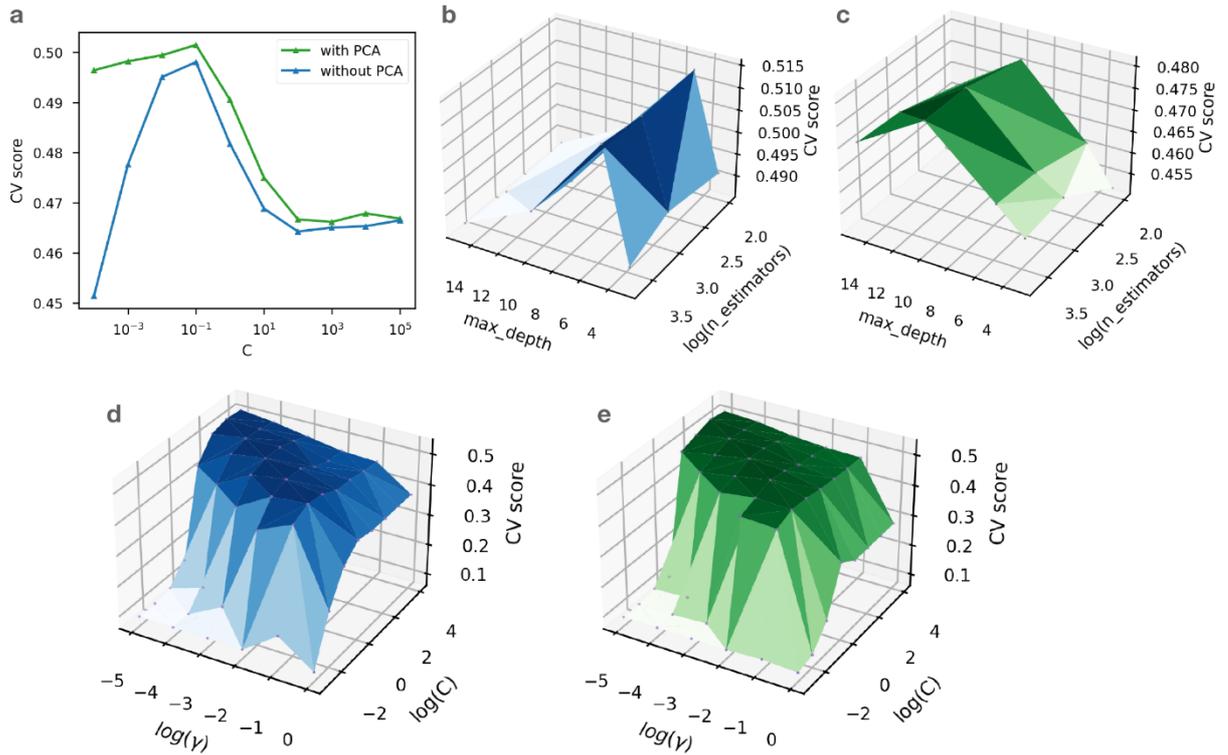


Figure S 3-3. Visualizations of the grid search using raw data. (a) LR, (b) RF without PCA, (c) RF with PCA, (d) SVC without PCA and (e) SVC with PCA.

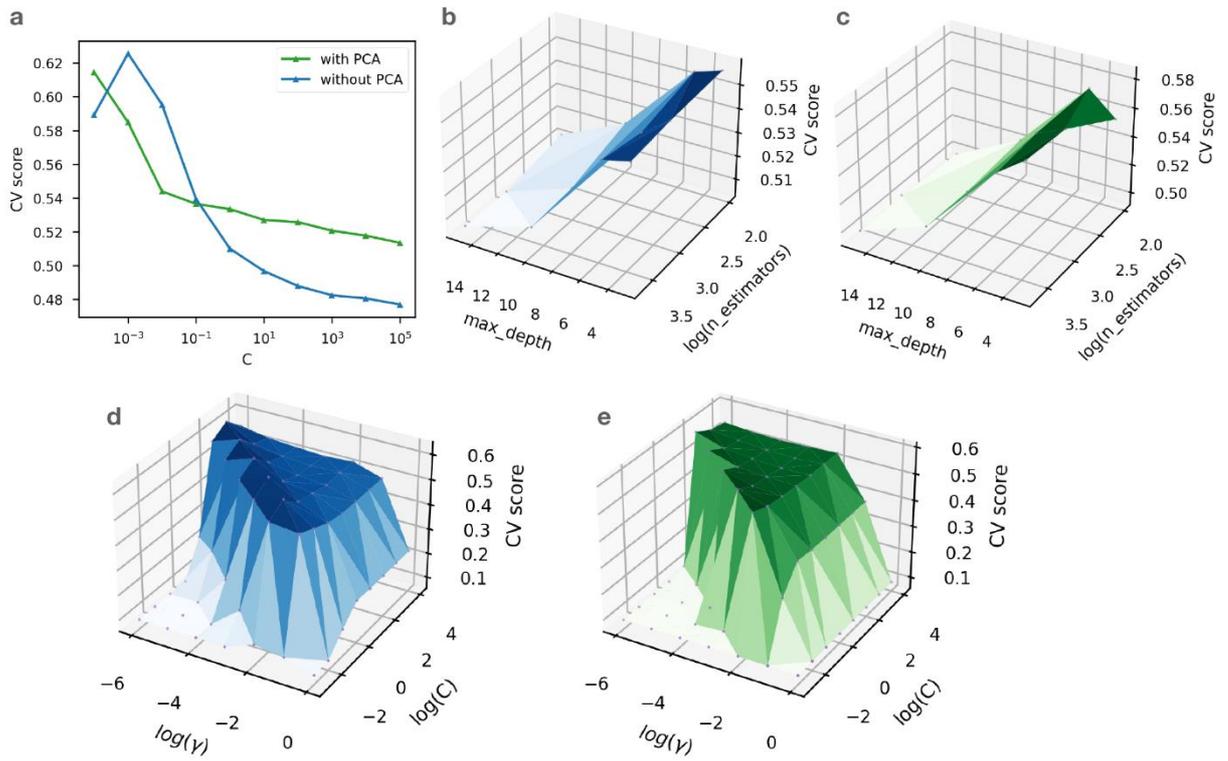


Figure S 3-4. Visualizations of grid searching using rolling data. (a) LR, (b) RF without PCA, (c) RF with PCA, (d) SVC without PCA and (e) SVC with PCA.

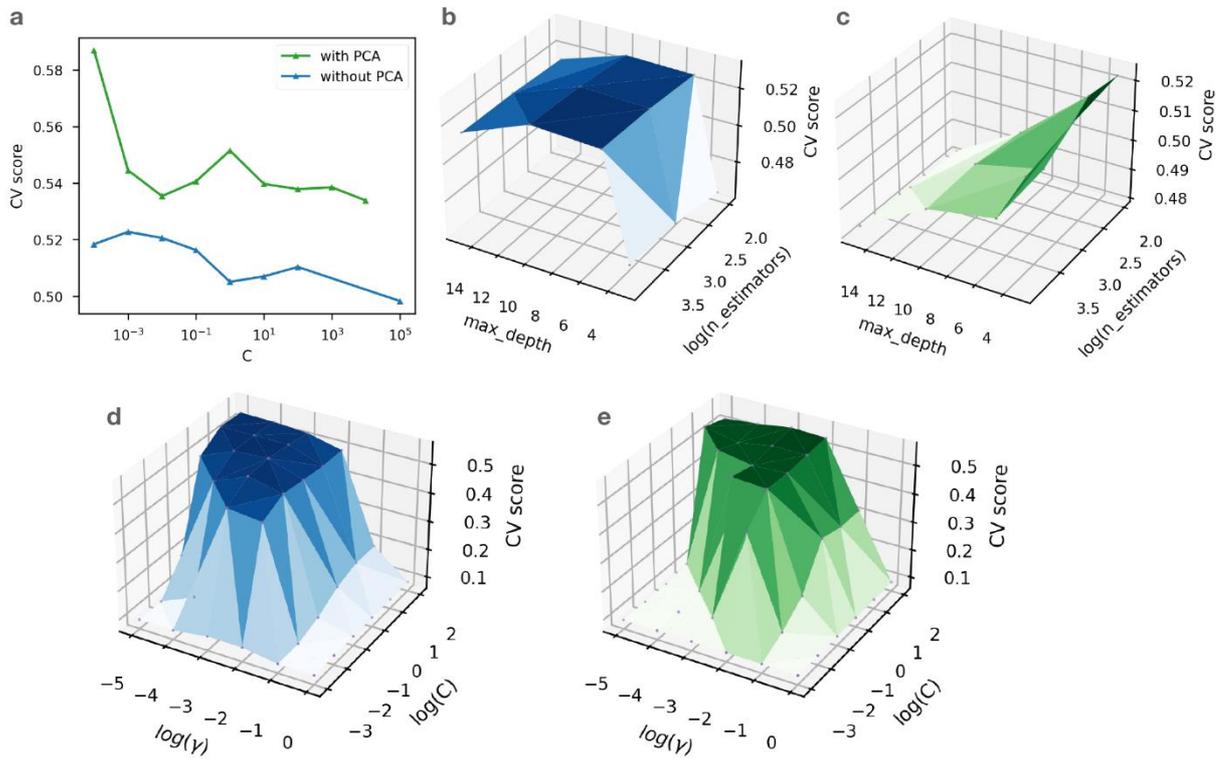


Figure S 3-5. Visualizations of grid searching using 2d data. (a) LR, (b) RF without PCA, (c) RF with PCA, (d) SVC without PCA and (e) SVC with PCA.

3.6.3. Test set results

The test set has 11 core sections. Their comparisons between sedimentologists' descriptions and predictions of the optimal model built using LR from the rolling data without PCA transformation are visualized in Figure S 3-6 and Figure S 3-7. Table S 3-2, stored in an excel file due to the large size, lists the error analysis result.

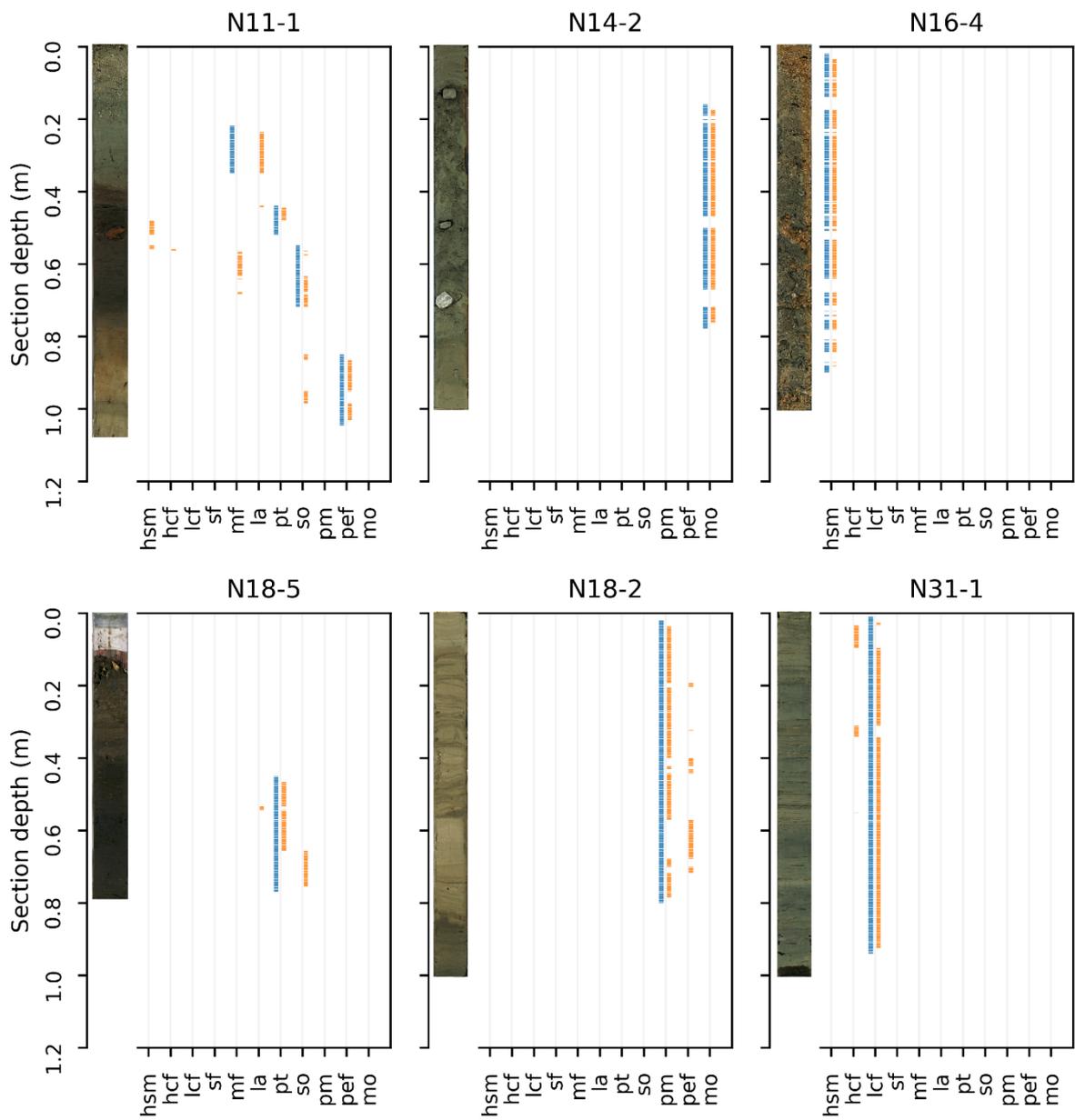


Figure S 3-6. Test set comparing sedimentologists' recognitions (blue) and classifications (orange) of the optimal model built using LR from the rolling data without PCA transformation. The core images are attached. For abbreviations refer to Table S 3-1.

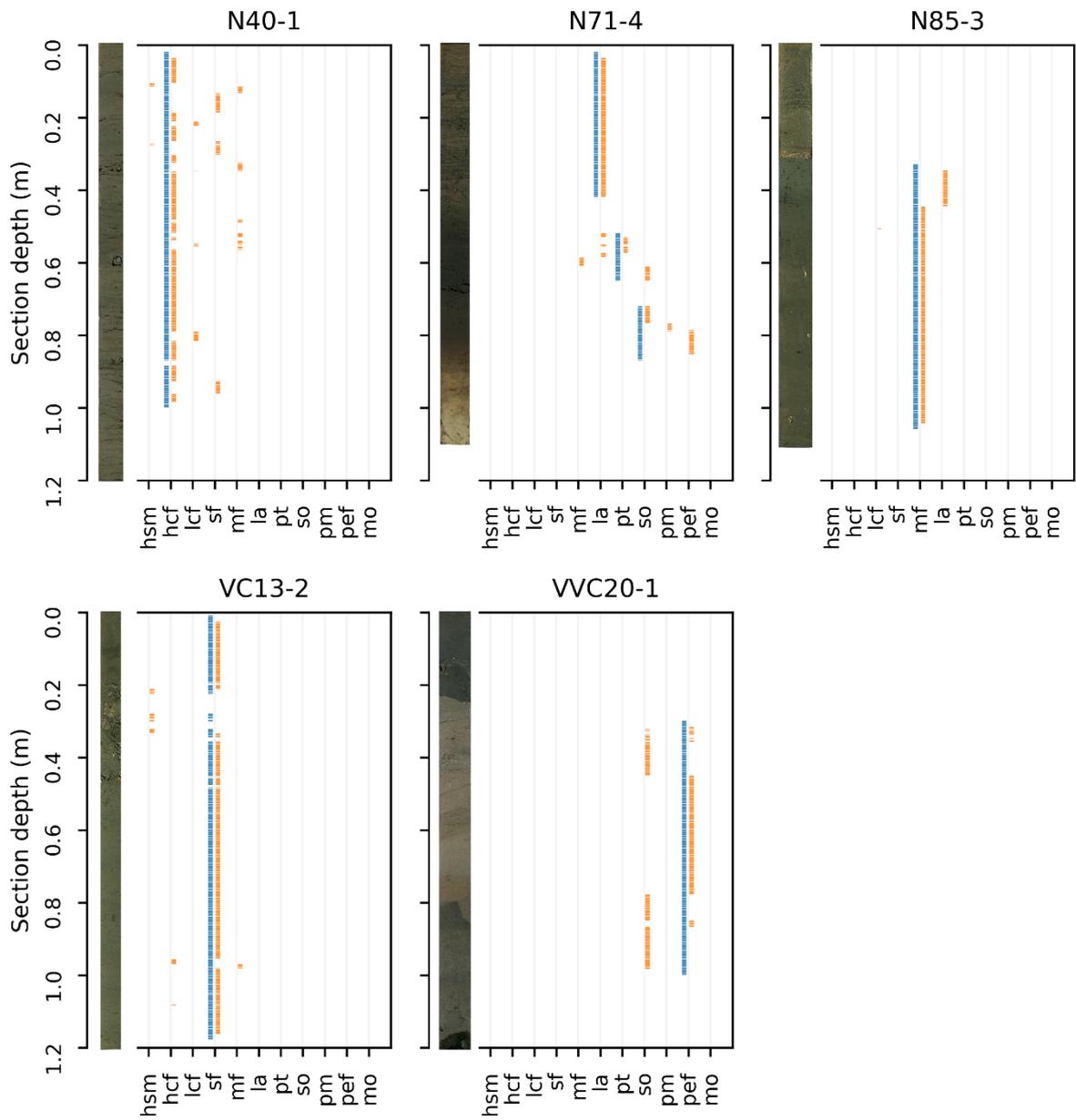


Figure S 3-7. Test set comparing sedimentologists' recognitions (blue) and classifications (orange) of the optimal model built using LR from the rolling data without PCA transformation. The core images are attached. For abbreviations refer to Table S 3-1.

Table S 3-2. Error analysis result: error_analysis_20220406.xlsx

4. Study III: Bulk chemistry quantification

Quantifying CaCO₃ and TOC in multi-regional marine sediments using XRF spectra - integrated application of geochemistry and machine learning

An-Sheng Lee ^{1,2} and Weng-Si Chao ³, Sofia Ya Hsuan Liou ², Ralf Tiedemann ³, Bernd Zolitschka ¹, Lester Lembke-Jene ³

¹ University of Bremen, Institute of Geography, Bremen, Germany

² National Taiwan University, Department of Geosciences and Research Center for Future Earth, Taipei, Taiwan

³ Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung, Bremerhaven, Germany

In submission

Scientific Reports

Author contributions

A. -S. L. and W. S. C.: Conceptualization, Methodology, Investigation, Data curation, Writing – original draft. These authors contribute equally.

R. T., L. L. J, B. Z., and S. Y. H. L.: Supervision, Writing – review and editing, Resources, Funding acquisition.

Abstract

Geochemical variations of sedimentary records contain vital information for studying Earth's history and to facilitate climate modeling and prediction. To obtain quantitative geochemical data in the laboratory is laborious, ultimately restricting their temporal and spatial resolution. Quantification based on fast-acquired and high-resolution X-ray fluorescence (XRF) core scanning data provides a solution to lift this restriction. Our study involves machine learning (ML) and powerful computing techniques to advance the quantifying progress. We target at calcium carbonate (CaCO_3) and total organic carbon (TOC) for quantification. The input data is raw XRF spectra instead of software-based extraction of elemental intensities to avoid bias and increase information. Our dataset consists of marine sediment samples retrieved from high- to mid-latitudes of the Pacific Ocean to extend the applicability of our quantification model from a core-specific to a multi-regional scale. The ML-built models are carefully evaluated in training and test sets as well as by a case study. The resulting models' performance (R^2 , CaCO_3 : 0.96 and TOC: 0.78) is estimated with uncertainty. This allows us to generate high-resolution bulk chemistry records without losing accuracy in the study area.

4.1. Introduction

4.1.1. Importance and limitations of bulk geochemistry

The oceans collect an extensive output from continental weathering. These dissolved elements and particulate products are eventually deposited on the seafloor due to biological and geochemical processes. Therefore, seafloor sediments record the history of continents and oceans. Qualitative and quantitative analyses of sedimentary components and their physical properties are fundamental to provide information for reconstructing paleoenvironmental changes of continents and oceans ⁹. For example, calcium carbonate (CaCO_3) as one of the most important biogenic components of pelagic marine sediments has drawn growing attention during the past decades in fields like paleoceanography and marine biology. Deep-sea carbonates act as the main sink and source of regulating the atmospheric CO_2 concentration ¹⁰. In order to continue interpretation and further investigations, laboratory analyses to obtain the principal data are essential. In this case, the determination of CaCO_3 content (wt%) is commonly calculated from the difference between total carbon content and total organic carbon content (TOC wt%; see Materials and methods). However, these laboratory analyses require sufficient time and labor. The temporal and spatial resolution is hence restricted. As an analog, a ten-meter-long marine sediment core normally takes 1 to 1.5 week to acquire bulk geochemistry measurements in 10 cm resolution.

4.1.2. Advantages of XRF core scanning and its quantification

With the advantages of high-resolution ($\geq 100\mu\text{m}$), non-destructive and rapid measurements, XRF core scanning techniques may lift the restriction of resolution. A ten-meter-long and opened marine sediment core can be scanned in 3 days to obtain elemental profiles with a 1 cm spatial resolution and replicates. This method has thus contributed significantly to systematically recording high-resolution geochemical profiles in sediments. Moreover, its applicability covers from soft sediments to speleothems, corals, rocks and tree sections¹⁵. Despite its advantages, the method has one major disadvantage: it only provides semi-quantitative measurements. The non-linear relation between quantitative and XRF measurements is caused by physical and matrix effects and a general lack of control on measurement geometry^{16, 115, 126}. As a result, quantification of XRF measurements has become a heated demand.

There are many stepstones on the path of quantifying XRF measurements. Several attempts to quantify elemental concentration from the XRF intensity via direct regression (i.e., Ordinary Squares Regression) have achieved moderate success (e.g.,^{16, 127, 128}). Weltje and Tjallingii (2008)⁷⁰ improved the quantification by introducing the Additive log-ratio and the Main Axis Regression to the workflow. The relation between elemental concentration and XRF elemental intensities is carefully dealt with by using detailed mathematic derivation (XRF spectrometry theory and statistical theory of compositional data) and empirical tests. However, the power of quantification is limited to “relative” element concentration, which is constrained by the input element assemblage. As a follow-up, Weltje et al. (2015)¹¹⁵ proposed a next level of improvement to increase the accuracy and quantify “absolute” elemental concentrations. They modified the previous workflow⁷⁰ by implementing Centred log-ratios and Partial Least Squares Regression.

The idea of covariance between matrix elements is hence included. Furthermore, it introduces Cross-validation (CV) to more rigorously evaluate the model's predictive power. To make this workflow easier to be applied, the software Xelerate was developed supported by a statistically robust sampling scheme (<http://www.mennobloemsma.nl/software.php>).

Based on thriving computing power and machine learning (ML) techniques^{22, 23}, a further stepstone of quantifying XRF measurements is reached. First, the application should not be limited to element concentration since there are other proxies, which are of interest for the scientific community and will benefit from high resolution, such as grain size, terrigenous input, opal, CaCO₃ and organic matter. Second, the elemental intensity to be quantified can be substituted by the emitted fluorescence energy and wavelength spectrum (in short: XRF spectrum, i.e., raw data), which records comprehensive information of scanned sediments but is a jump of two orders of magnitude in data dimension. The commonly used elemental intensity is generated after by processing the XRF spectrum. Fine-tuning of the related software settings needs care and experience. To produce semi-quantitative data rapidly, this procedure often receives less attention than necessary. Consequently, using the XRF spectra may avoid this manual bias, especially when dealing with a high quantity of sediment cores. Third, the data space of logarithm can be extended to an infinite non-linearity space efficiently by a mathematical trick, kernel function¹¹⁷. It is expected to give learning algorithms higher capability to cope with the non-linear relation between semi-quantitative XRF and desired quantitative elemental measurements. Fourth, the application of a model was often considered as site- or even core-specific. Having a larger quantity data from a number of cores, a kernelized

learning algorithm may lift this limitation to build a large regional quantification model, which smoothens the process of acquiring high-resolution quantitative data.

Herein, we propose an approach of building models quantifying two commonly used geochemical proxies (CaCO_3 and TOC contents) from XRF spectra and involving ML techniques and high-performance computing, by using marine sediment cores that cover multiple regions from the northern and southern pelagic Pacific Ocean (Figure 4-1). Two preprocessing algorithms, Principal Component Analysis (PCA) and Non-negative Matrix Factorization (NMF), and three supervised ML algorithms, Ridge Linear Regression (LR), kernel Support Vector Machine (SVM) and Random Forest (RF), are included for searching the optimal model. PCA and NMF allow the extraction of essential information for facilitating the following learning process. SVM and RF are the algorithms capable of learning non-linear relation, while LR is a linear algorithm for comparison. The workflow is schematically shown in Figure 4-2. Hopefully, it will provide kick-off opportunities to include more measurements to cover extensive application of quantification, such as different proxies' measurements and scanner types. Meanwhile, all the executive codes are open source. The users can easily adopt and modify the codes for their own study, rather than being restricted to a certain software.

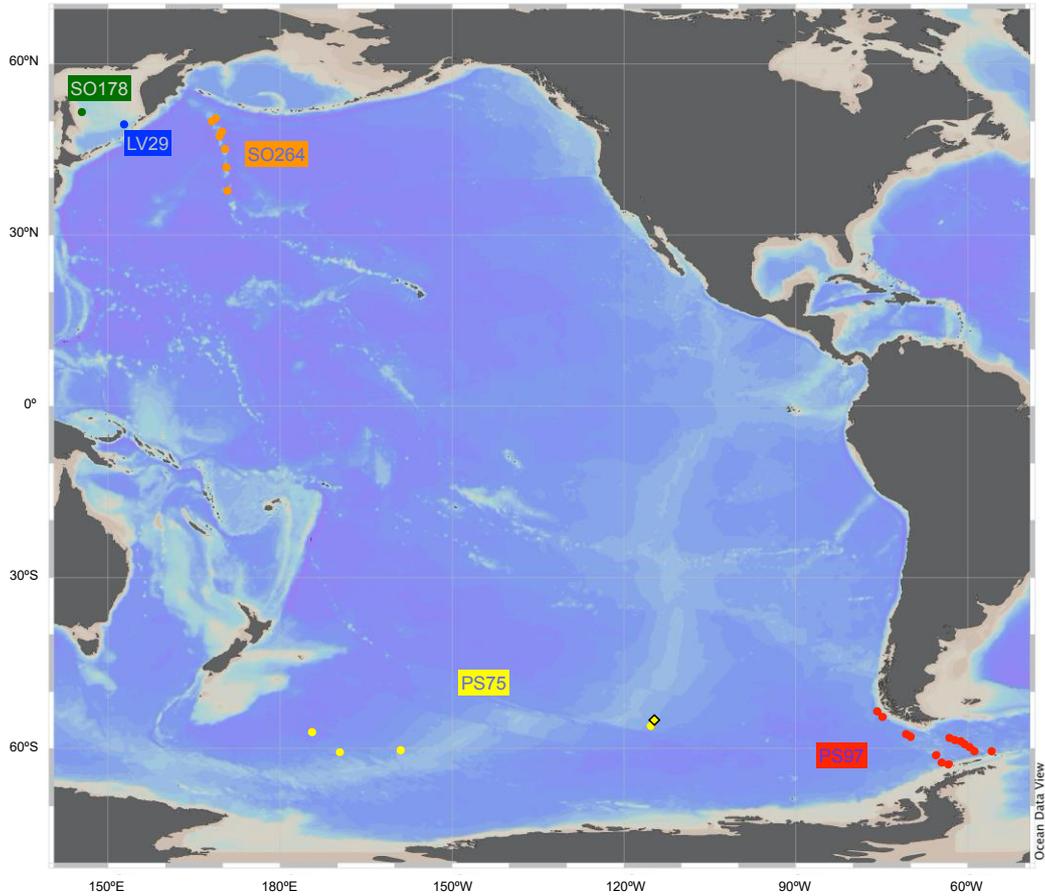


Figure 4-1. Map of selected pelagic sediment cores, marked by colored dots related to individual labelled cruises. The yellow rhombus represents the case study core PS75/056-1.

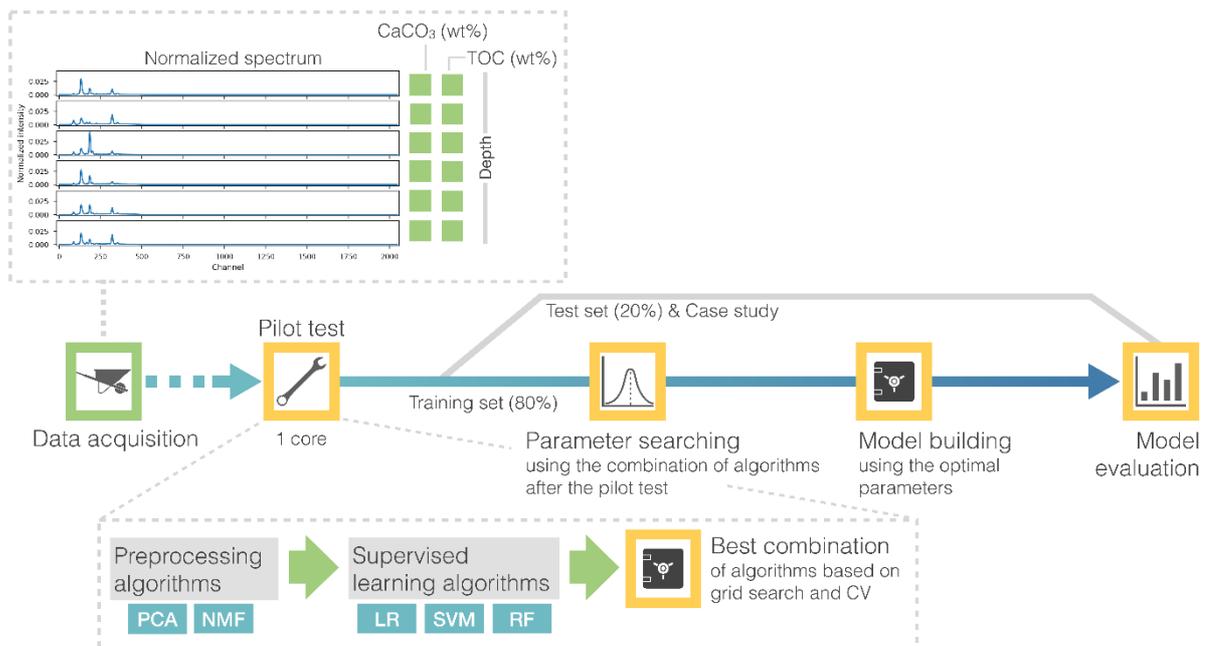


Figure 4-2. Workflow for the approach.

4.2. Results

4.2.1. Determination of optimal models

The pilot test shows the workflow using SVM to learn from the NMF-transformed data provides the most promising results. The linear algorithm (LR) has disadvantages in giving good predictions compared to the SVM workflow. Consequently, this workflow was carried out for the training set. After grid searching for the parameters, the optimal models were built using the settings listed in Table 4-1. Details of the grid search results are provided with Supplementary Material.

Table 4-1. Parameter settings and scores for the optimal model of the analytes calcium carbonate (CaCO_3) and total organic carbon (TOC). $n_components$ is for NMF. C and γ are for SVM. Cross-validation (CV) score stands for the learning performance in the training set. N describes the data amount in each subset. R^2 , 95% confidence interval (CI) and root-mean-square error (RMSE) estimate the performance and uncertainty of the optimal models in the test set and case study (core PS75/056-1). CI and RMSE both have unit as wt%. *Negative R^2 is considered as 0.

Analyte	$n_components$	C	γ	CV score (R^2)	Test set		Case study	
					R^2	95% CI	R^2	RMSE
CaCO_3	4	100	1000	0.87	0.96	[-7.04, 7.00]	0.82	6.66
TOC	13	10	1000	0.79	0.78	[-0.13, 0.13]	0*	0.06

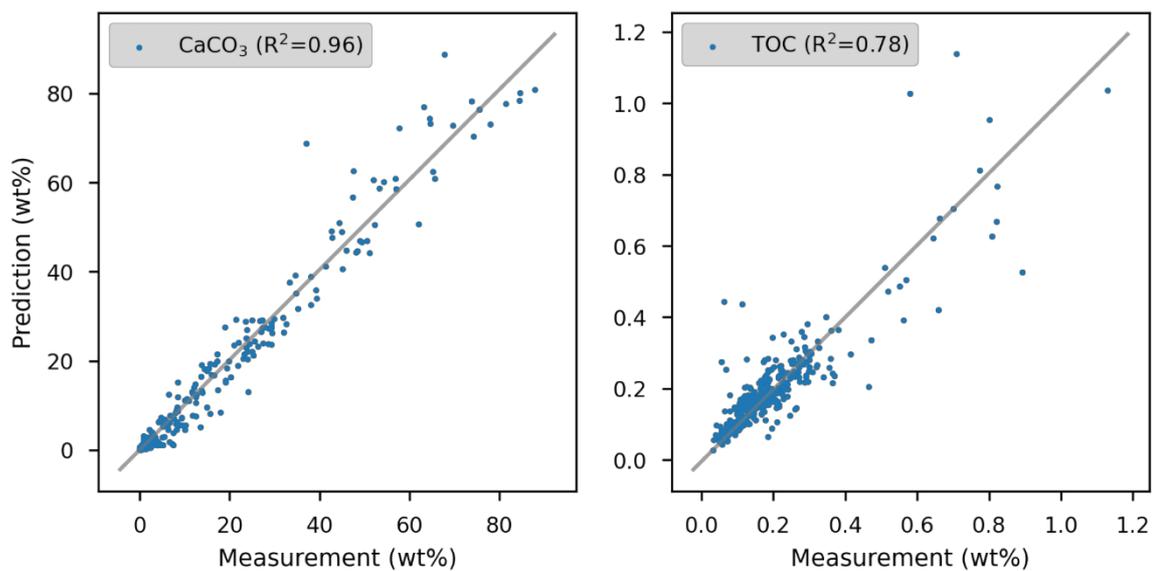
4.2.2. Evaluations of test set and case study

As approaches tend to enhance the accuracy of the model in the training data only, this causes a common mistake while evaluating a built model. For example, there are 20 CaCO_3 and 100 XRF measurements in a core. Operators often try to find a regression between these 20 CaCO_3 and corresponding XRF measurements giving the highest R^2 . They hence overestimate the regression's accuracy outside of the training data. This is called overfitting in ML tasks ¹¹⁷. The model's generalization beyond the training data should be equally important.

Our optimal model accuracies in the training set is evaluated by CV scores. Then, they were evaluated again in the test set. Table 4-1 shows our models have delicate accuracy in both training and test sets as represented by R^2 . The test scores are equal or even higher compared to the CV scores, which excludes the effects of overfitting during training. The 95% confidence intervals estimate narrow and constrained uncertainties of the models. The goodness of fitting is visualized in Figure 4-3 (a).

The measurements and predictions with uncertainty in the case study are shown as a whole core variation example in Figure 4-3 (b). Some obvious errors appear in Figure 4-3 (b), but mostly they stay within the confidence interval. The models' R^2 in the case study drops while comparing to the test set (Table 4-1). Yet, the root-mean-square errors (RMSE) remain in low values (CaCO_3 : 6.66 wt% and TOC: 0.06 wt%), which are within the confidence interval calculated from the test set.

a. Test set



b. Case study

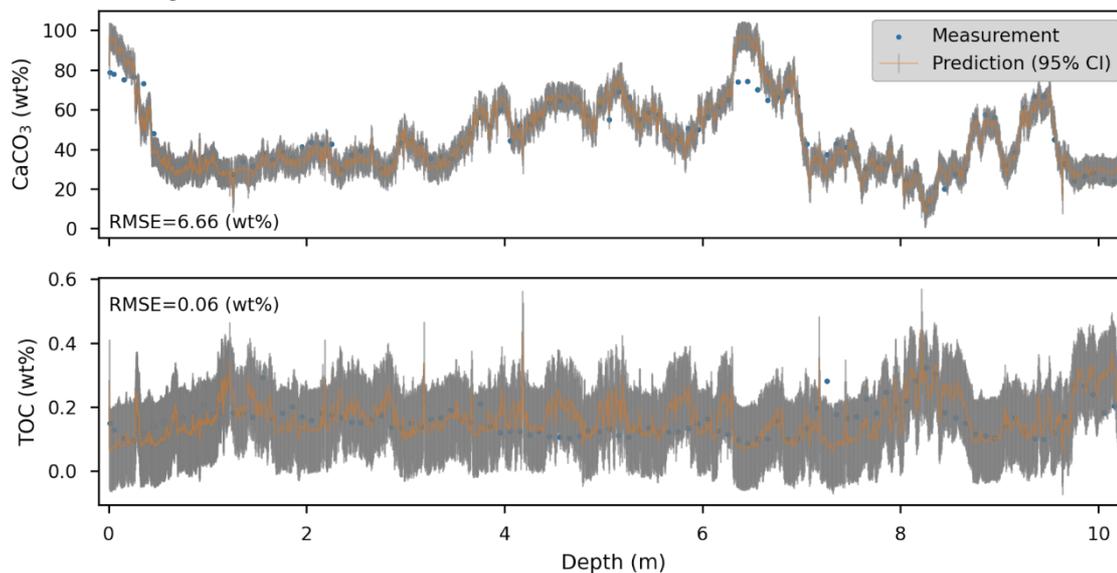


Figure 4-3. Measured versus predicted CaCO_3 and TOC contents in (a) the test set and (b) the case study of core PS75/056-1. The predictions are displayed with 95% confidence interval, which is shown in Table 4-1.

4.2.3. Outcome: quantified data of CaCO₃ and TOC

As the optimal models have satisfactory accuracy during evaluation, they were implemented to quantify the whole dataset of cores. The total number of data points with bulk geochemistry is increased from <2,000 to 57,240. As a result, for the core intervals which have not sampled for bulk geochemistry now have quantified values. All these high-resolution (1 cm) quantified CaCO₃ and TOC contents in the study area are available in Supplementary Material. In the whole quantified data, only 410 data points (0.72 %) are over 100 wt% and no negative value. Figure 4-3 (b) illustrates the improved resolution of bulk chemistry.

4.3. Discussion

There are two possible reasons to explain the good performance of our approach. First, the direct use of XRF spectra instead of elemental intensities after software processing helps the ML algorithms to learn the information without any bias caused by manual fine-tuning. This fine-tuning of software settings needs sufficient experience to adjust for sediment-property changes. For instance, the change of water content or organic matter acquires corresponding adjustments of the settings. However, these adjustments are commonly ignored due to the lack of experience or tedious labor when encountering long records, which leads to biased elemental intensities. Secondly, compared to software-processed elemental intensities or element ratios, the spectrum provides more hidden information recorded in the whole fluorescence signals (e.g., water content, organic matter). With the help of increasing computing power and non-linear ML techniques^{22, 23}, this information should no longer be waived. As a result, after learning from a certain amount of data points covering the entire variation of sediment properties,

our ML-built models are able to determine unbiased multi-elemental and matrix information and capable of quantifying the bulk measurements.

Although the R^2 of TOC in the case study is actually off when compared to the test set, it does not imply the degraded performance. The denominator of R^2 is the total sum of squares, which makes R^2 sensitive to the variance of the actual measurements. Hence, the dropping R^2 is misled by the limited variance in the case study (0.004) comparing to the variance in the test set (0.019) while its numerator (the sum of squares of residuals) remains stable. The steady RMSE in the case study proves this explanation. In fact, 0.06 wt% of TOC might actually below the detection limit of quantitative measurements.

Unlike conventional quantification methods (e.g., ^{16, 115}e.g., ^{16, 115}e.g., ^{16, 115}e.g., ^{13, 112}e.g., ^{13, 112}e.g., ^{13, 109}e.g., ^{13, 109}e.g., ^{13, 109}e.g., ^{13, 99}e.g., ^{13, 99}e.g., ^{11, 95}e.g., ^{11, 95}), which build site- or core-specific models, our approach shows that general models can be constructed for quantifying CaCO_3 and TOC to multi-region synchronously. Our models have the ability of generalization, i.e., they are able to be implemented on cores other than those with bulk chemistry measurements. Thus, the demand of repeated building model is eliminated. Eventually, our models would be able to apply on the rest of the cores from the specified cruises which cover multiple sectors of Pacific Ocean. Our demand of conventional laboratory measurements is reduced noticeably. In this study, the laboratory measurements take only 3.5% of the total data amount (i.e., measurements in 1 cm interval). The needed amount laboratory measurements for building quantification model is greatly reduced.

4.4. Guidelines for applications

The concept of machine learning is data-driven, which means extrapolations might be risky for the models. Although SVM compared to tree-based algorithms (e.g., Decision Tree and RF) does not severely suffer from this issue and our data coverage is not small, it is better to be cautious²¹. Another notable point is the boundary conditions of the data including the sediment type, sediment age range, XRF core scanner, X-ray tube and scanning settings while directly applying our models. Any changes of these categories might result different XRF spectra and lead to a mistaken prediction. An efficient way of examining boundary conditions is making sure that XRF spectra are produced with the same setting and scanner fall into our spectral range, which is listed in Supplementary Material. The rest boundary conditions of our dataset are stated in Materials and methods section for future applications.

For studies retrieving sediment cores in nearby regions: our models are ready to use as long as the cores are scanned by the same type of XRF core scanner, X-ray tube and settings. No laboratory measurement and high-performance computing is necessary. A common personal computer or laptop with capability of basic Python coding is sufficient for adopting our models. For the studies having cores far from our study area or using different core scanners: our workflow to build models is suitable for developing own models. The models and the developing codes are all opened on Github (https://github.com/dispink/CaCO3_NWP).

4.5. Conclusions

This approach successfully builds models that can quantify bulk chemistry directly from XRF spectra. The usage of XRF spectra reduces manual bias and increases input information for the ML algorithms. The wide data coverage and the power of ML and computing techniques lift the model's limitation off the site-specific scale. Moreover, quantification of CaCO_3 and TOC contents is novel. Our optimal models' generalization is guaranteed by carefully evaluating training and test sets and the case study. The uncertainty of predictions is estimated in 95% confidence interval from the test set.

All the retrieved cores from the northern and southern pelagic Pacific Ocean now could have quantified high-resolution (1 cm) bulk measurements, which can be adopted for further investigation like mass-balance and flux calculation. With enhanced resolution, the variation gives us possibilities to interpret more details and features from the records.

Guidelines are provided for future users. To make our multi-region models thoroughly inclusive, we are eager to incorporate more measurements, such as different core locations, scanner types and settings. Hopefully, this approach brings future studies a step forward to retrieve high-resolution quantitative bulk chemistry without losing accuracy.

4.6. Materials and methods

4.6.1. Sediment cores and bulk measurements

In this study, spectra of 27 XRF-scanned marine sediment cores, including published and unpublished records, together with their corresponding bulk chemistry measurements were examined (in total, 2285 TOC samples and 2222 TC samples; Supplementary Material). The investigated cores are mostly retrieved from two expeditions: cruise SO264 in the subarctic Northwest Pacific with R/V SONNE in 2018 and cruise PS97 in the central Drake Passage with RV Polarstern in 2016. Another four cores and the case study core were recovered in the Pacific sector of the Southern Ocean during PS75 in 2009/2010. Further two cores were recovered in the Okhotsk Sea during cruise KOMEX II with R/V Akademik Lavrentyev in 1998 and cruise SO178 in 2004. The cover area of all investigated cores is mainly spreading over the high- to mid-latitude Northwest Pacific (37°N-52°N) and the Pacific sector of the Southern Ocean (53°S-63°S), with a water depth coverage from 1211 m to 4853 m.

The investigated cores are dominated by pelagic sediments and mainly consist of calcareous and siliceous ooze, and non-carbonaceous fine-grained sediments. We use cores from the pelagic ocean instead of cores close to the shore to avoid influences of coastal erosion and fluvial input (organic matter, terrigenous input). The age range of the majority of the sediments spanned from Holocene to mid-Pleistocene¹²⁹⁻¹³⁵.

To determine bulk sediment parameters, immediately after opening of the sediment cores on board, we took syringe samples of 10 cubic centimeters in 10 cm intervals from the working halves of the core into pre-weighed glass vials. All samples were stored at 4° C until further shore-based processing. The samples in this study were weighed before and after freeze-drying, homogenized and measured for TC and TOC by

a CNS-analyzer (Elementar vario EL III) and a carbon/sulfur analyzer (Eltra CS-800) in the laboratories at AWI in Bremerhaven. The CaCO_3 content was then calculated from the difference between TC and TOC as follows:

$$\text{CaCO}_3 \text{ wt\%} = (\text{TC} - \text{TOC}) \times 8.333 \times 100\%$$

4.6.2. Setting of the Avaatech XRF scanning and data compilation

All XRF-scanning measurements were carried out with an Avaatech XRF core scanner at the AWI in Bremerhaven. The X-ray excitation scanning settings were 10 kV at 150 mA with no filter for a count time of 10 s, and a rhodium target X-ray tube was deployed. Sample distance was 10 mm, with a 10x12 mm slit size. In order to include comprehensive information and avoid the software bias, which could be caused by manual operation, the raw spectrum (2048 channels) for each scanning point was chosen instead of processed elemental intensities. The XRF spectra were normalized by the sum of channel signals for each data point to prevent bias from X-ray tube aging. The bulk geochemistry (CaCO_3 , TOC) were collected and nature-logarithm transformed to avoid predicting negative contents later in the model. The XRF spectra were aligned to the bulk chemistry measurements by depth.

4.6.3. Pilot test

Quantifying the XRF spectra to bulk geochemistry is a regression task carried out by ML algorithms to learn the relation between both data. To find a suitable combination for preprocessing and supervised learning algorithms, the XRF spectra and CaCO₃ content from one of the recently retrieved cores (SO264-15-2) were selected as pilot data (data amount: 38). We tested two preprocessing algorithms (Principal Component Analysis - PCA and Non-negative Matrix Factorization - NMF) and three supervised learning algorithms (Ridge Linear Regression – LR, kernel Support Vector Machine - SVM and Random Forest - RF). PCA and NMF are commonly applied for source separation, extracting vital information from the data ^{27, 28, 136}. SVM has abilities to explore data relations in the infinite non-linear space and tolerate data noise ¹²⁰. RF provides a well-regularized learning power in non-linear space based on its tree-based design ^{119, 137}. LR is an L2 regularized Ordinary Least Squares regression ¹¹⁷, included as a reference for the performance of the linear algorithm.

The `n_component` parameter for NMF was initially set following the PCA's result (explained variances of principal components). The parameters (alpha for LR, C and gamma for SVM, `max_depth` and `n_estimators` for RF) control the regularization level of the algorithms, which affect under- and over-fitting issues ^{21, 73}. There is no universally suitable value for them, so we intuitively grid searched parameters and found a set of optimal parameters, which builds a model with best performance. Since the data is not noise-free, this grid search was integrated with 5-fold cross-validation (CV) ¹¹⁷ to be robust. The score (performance) is presented by an averaged coefficient of determination (R^2) in CV.

4.6.4. Model training and evaluation

After specifying the workflow with the best score for the pilot test, the training set, split from the whole dataset (random 80% of data points, Figure 4-4), was used to train our model. The grid search was implemented again for finding optimal parameters, but integrated with 10-fold CV to increase the score's statistic robustness. The final models for quantifying CaCO₃ and TOC were subsequently built by adopting the optimal parameters. Details of the grid search strategy are provided by Supplementary Material I.

Evaluating their performance in a pristine data subset is essential to understand the models' generalization, i.e., how good our models are in quantifying bulk chemistry from other cores. The test set comprises the data left after partitioning the training set from the whole dataset (Figure 4-4). Since it is a set of random data, it cannot illustrate a comparison between measurement and prediction in a manner of whole core with uncertainty. This study also preserved a core of data, as a case study (PS75/056-1, Figure 4-4), to evaluate the built models' performance. During evaluation, the models read the XRF spectra of this subset of data to predict CaCO₃ and TOC contents. The performance was calculated by R² and root mean squared error (RMSE) between actual and predicted values. Meanwhile, errors in the test set were used to construct 95% confidence interval based on the t-distribution. The uncertainty of our models was thus estimated.

All computations and visualizations were conducted using the SciPy ecosystem in Python ^{73, 79, 80, 83, 124, 125}.

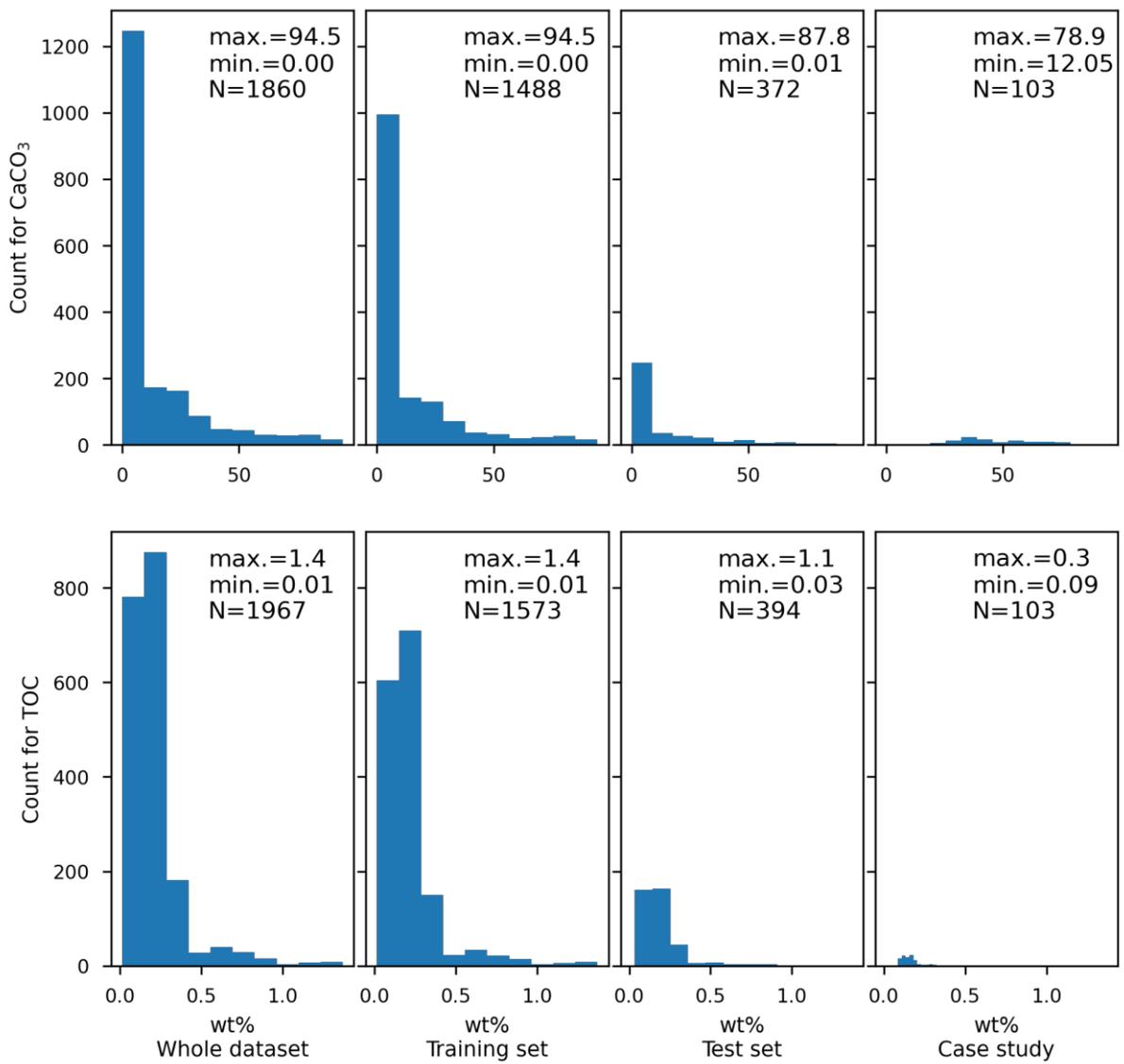


Figure 4-4. Data distributions of CaCO_3 and TOC for the whole dataset, training set, test set and case study.

4.7. Acknowledgments

We thank the crew and the science parties for their contributions to all expeditions as well as technicians and students at the AWI Marine Geology Laboratory in Bremerhaven who contributed effort and time to all the samples and measurements. We also acknowledge Ben Marzeion and Timo Rothenpieler (Klimageographie, University of Bremen) for generously supporting their high-performance computing power.

We acknowledge financial support from the Ministry of Science and Technology of Taiwan (project number: MOST 110-2116-M-002-023), the AWI institutional budget under the “PACES-II” (2018-20), the German Federal Ministry for Education and Research (BMBF) through grants no. 03G0264B “SO264-EMPEROR” and 3F0785B “NOPAWAC”.

4.8. Supplementary materials

4.8.1. Detailed grid search results

During the pilot test, the number of components was chosen at 5 before the explained variance flattened (Figure S 4-1-(Left)). This number was also adopted later for the initial grid searching of Non-negative Matrix Factorization (NMF). Figure S 4-1(Right) indicates that the combination of NMF and kernel Support Vector Machine (SVM) gives the best performance ($R^2 = 0.91$). The Ridge Linear Regression (LR) shows the slightly worse performance, while the Random Forest (RF) has the worst performance.

There are three parameters need to be fine-tuned in building models, $n_components$ for NMF, C and γ for SVM. Based on the Principal Component Analysis (PCA) result of the pilot test (Figure S 4-1), we initially tried $n_components$ from 5. In each setting of component, ranges of C and γ have been searched in logarithm to find the model having best performance. When the optimal combination of parameters locates away from the edge of plot (Figure S 4-2 and Figure S 4-3, i.e., ranges of parameters), it is chosen to build the optimal model. The optimal component amount for TOC model was set to 13, at the edge of searching range. The reason why not kept searching upwards is that the model's fitting time extends noticeably after increasing it.

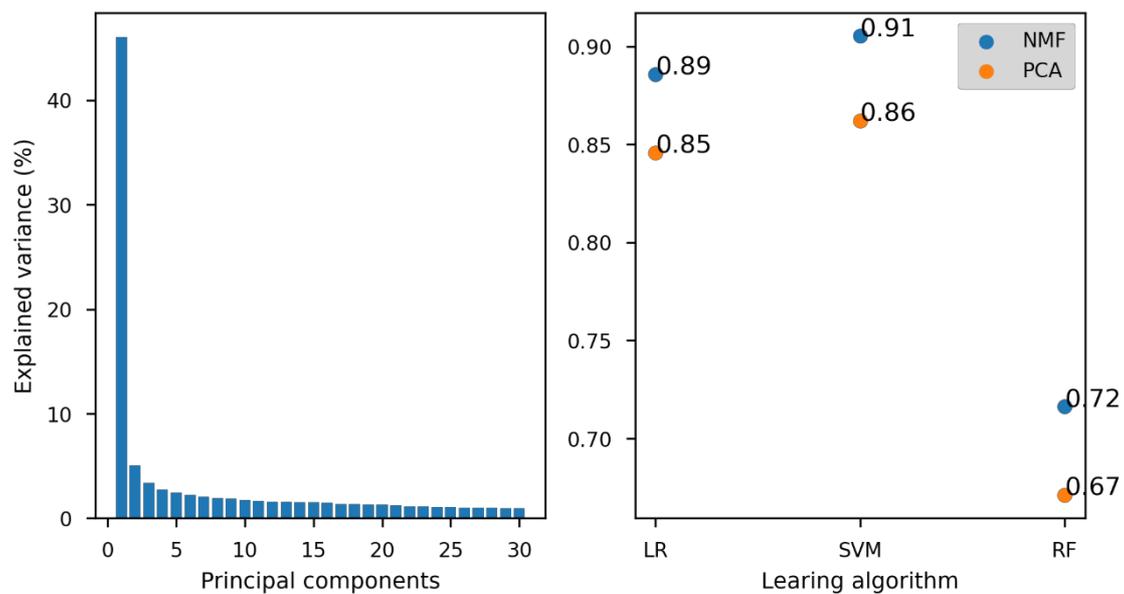


Figure S 4-1. Results in the pilot test. Left: Explained variance of each principal component. Right: The best performance (R^2) of each workflow.

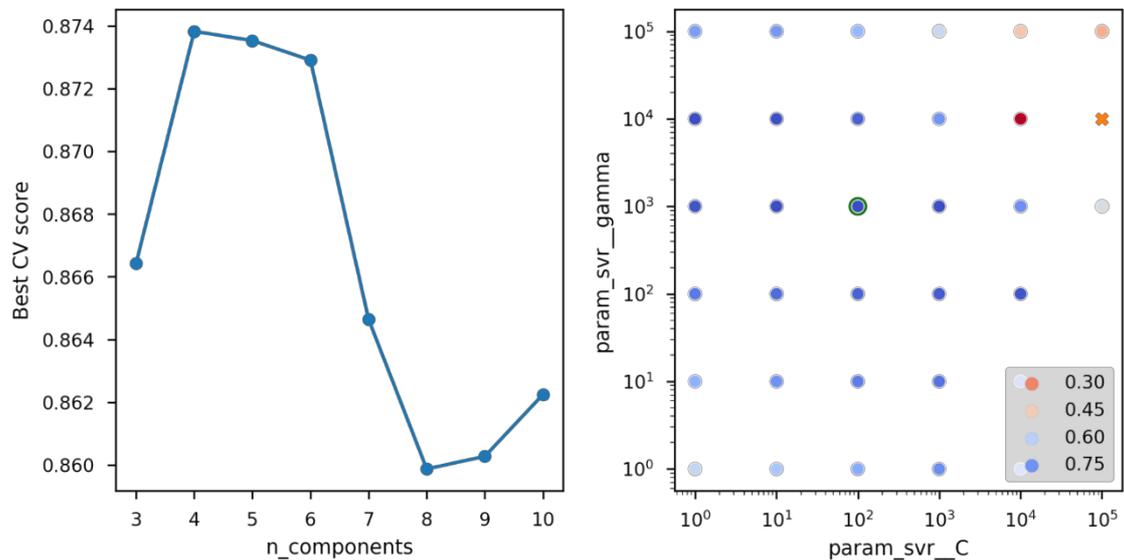


Figure S 4-2. Visualization of CaCO_3e 's grid search results. Left: Best CV score in each setting of $n_{\text{components}}$ for NMF. Right: CV score grid of the pair of SVM parameters under $n_{\text{components}} = 4$. The optimal pair of SVM parameters is marked by green circle.

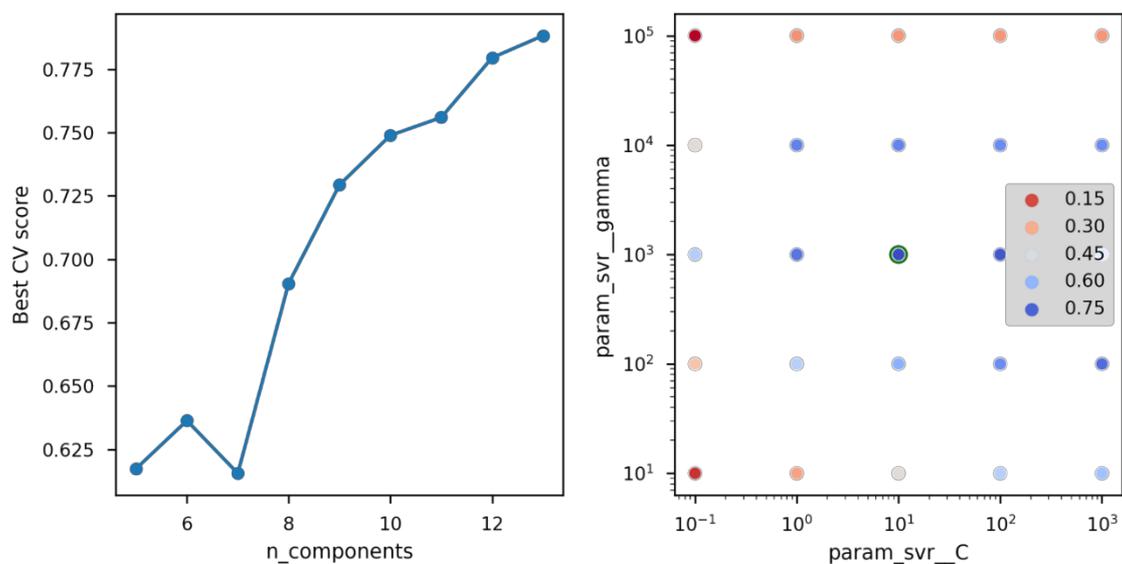


Figure S 4-3. Visualization of TOC's grid search results. Left: Best CV score in each setting of $n_components$ for NMF. Right: CV score grid of the pair of SVM parameters under $n_components = 13$. The optimal pair of SVM parameters is marked by green circle.

Dataset S1. Grid search results in the pilot research:

pilot_grid_ridge_20210823.csv

pilot_grid_svr_20210318.csv

pilot_grid_rf_20210822.csv

Dataset S2. Grid search result of building models for CaCO₃ and TOC:

grid_caco3+toc_compile_20210823.csv

4.8.2. Info and exported data

Dataset S3. High-resolution quantified CaCO₃ and TOC: predict_20220104.csv

XRF setting: Avaatech XRF core scanner with the X-ray excitation scanning settings, 10 kV at 150 mA with no filter for a count time of 10 s, and a rhodium target X-ray tube was deployed.

Dataset S4. Spectral range of our dataset: spectral_range_20220104.csv

Supplementary Table. Station list of the studied cores for the whole dataset, and the relevant information including amount of measurements, dating methods, core age and reference: supplementary table_station list and info.xlsx

5. Concluding remarks and outlook

5.1. Trials of core scanning techniques

This work selected representative core sections in the Wadden Sea area to demonstrate data analysis using unsupervised ML (PCA) on multiple core scanning data (MS, CT, XRF, digital photography). A novel data visualization, density plot, was presented to characterize these scanning data in 8 ‘a priori’ recognized sediment facies (moraine, eolian/fluviol, soil, peat, lagoonal, sand flat, channel fill, and beach-foreshore). The result shows these sediment facies can be differentiated by their physical and chemical properties from the scanning data, which confirms that the human-made lithological descriptions can be reproduced by time- and cost-efficient (semi)quantitative and high-resolution digital measurements. Therefore, it gives the possibility to implement ML for further automatic facies classification using scanning data.

5.2. Automatic sediment facies classification

As a follow-up approach, more sediment core sections and facies were included. A series of feature engineering and supervised ML algorithms were implemented to develop an optimal model that can automatically classify sediment facies by reading elemental profiles acquired from the XRF core scanning. A simple but powerful feature engineering that elevates the efficiency of ML algorithm and simulates the behavior of sedimentologists is proposed. The model’s performance was carefully evaluated by involving domain knowledge (i.e., sedimentology). It has a promising accuracy (78%) and a minor issue of fragmentation. The restrictions of small data dimensions and target class variety were overcome in this approach by involving 20k data points and 11 sediment facies varying from late Pleistocene terrestrial deposits to Holocene shallow

marine sediments. Furthermore, the model offers a valuable ability, highlighting critical sediment sections for sedimentologists. This would increase the classification capacity by redistributing resources more efficiently.

Even though the errors made by the model are primarily unavoidable bias, we can still optimize our model to be more human by better simulation of sedimentologists' behavior. With a flexible observational window and a bigger picture of environmental interpretations in mind, neural network algorithms' complex or customized architecture, like long short-term memory, may cause improvements. In addition, more kinds of scanning data, like digital images and hyperspectral imaging, can be included for comparing performance. The result is expected to be a stepstone for developing a commercialized automation protocol including both the front-end data acquisition and back-end classification in the near future.

5.3. Bulk chemistry quantification

Besides the classification problem, the thesis presents another field of ML application, regression problem. High resolution of CaCO_3 and TOC variations in deep-sea sediments can now be acquired with less labor cost via this approach. The direct usage of XRF spectra to quantify these two bulk chemistry measurements reduces manual bias during software tuning. It also increases input information for the ML algorithms, which results in promising accuracies in the test set (R^2 of CaCO_3 : 0.96, TOC: 0.78). The degradation of TOC model's R^2 in the case study is misled by the extremely small variance of TOC values in the case study. When considering the RMSE, the model's performance is still acting within the 95% CI calculated from the test set. In addition, the broad data coverage (multiple regions from the high latitude of the Pacific Ocean) and

the power of ML techniques lift the model's limitation off the site-specific scale that commonly happens in previous studies.

In the future, the power of quantification will be escalated by extending the data coverage in sediment type and core scanner settings. The information of these variances also needs to be included in the input data to assist machine learning. Moreover, the target class can be enriched to other useful proxies, such as diatom, opal contents and grain size, which we believe are in heavy demand.

In summary, the hypothesis proposed in the beginning is proved by including a more comprehensive database. The performance of automation empowered by ML is therefore enhanced by overcoming the limitation of small data coverage and applicability. Rigorous evaluation is simultaneously achieved to guarantee the progress. This thesis provides generalizable and fully open-source methodological blueprints for the presented works in order to facilitate the ML application in various geoscience subjects. The sustained advance of our capability for studying the Earth is then guaranteed. More groundbreaking discoveries will fuel the evolutions of science and civil community.

6. References

1. Reineck, H.-E., Singh, I. B. *Depositional Sedimentary Environments: With Reference to Terrigenous Clastics*. (Springer Berlin Heidelberg, 1986).
2. Zolitschka, B., Lee, A.-S., Bermúdez, D. P., Giesecke, T. Environmental variability at the margin of the South American monsoon system recorded by a high-resolution sediment record from Lagoa Dourada (South Brazil). *Quaternary Science Reviews* **272**, 107204 (2021).
3. Streif, H. Sedimentary record of Pleistocene and Holocene marine inundations along the North Sea coast of Lower Saxony, Germany. *Quaternary International* **112**, 3-28 (2004).
4. Leng, M. J., Marshall, J. D. Palaeoclimate interpretation of stable isotope data from lake sediment archives. *Quaternary Science Reviews* **23**, 811-831 (2004).
5. Whitehouse, R. J. S., Harris, J. M., Sutherland, J., Rees, J. The nature of scour development and scour protection at offshore windfarm foundations. *Marine Pollution Bulletin* **62**, 73-88 (2011).
6. Kropelin, S., *et al.* Climate-driven ecosystem succession in the Sahara: The past 6000 years. *Science* **320**, 765-768 (2008).
7. Fujii, T., *et al.* Geological setting and characterization of a methane hydrate reservoir distributed at the first offshore production test site on the Daini-Atsumi Knoll in the eastern Nankai Trough, Japan. *Marine and Petroleum Geology* **66**, 310-322 (2015).
8. Coughlan, M., Long, M., Doherty, P. Geological and geotechnical constraints in the Irish Sea for offshore renewable energy. *Journal of Maps* **16**, 420-431 (2020).
9. Schulz, H., Zabel, M. *Marine Geochemistry*, 2nd edn. (Springer, 2006).
10. Archer, D., Maier-Reimer, E. Effect of deep-sea sedimentary calcite preservation on atmospheric CO₂ concentration. *Nature* **367**, 260-263 (1994).
11. Ross, P. S., Bourke, A., Fresia, B. A multi-sensor logger for rock cores: Methodology and preliminary results from the Matagami mining camp, Canada. *Ore Geology Reviews* **53**, 93-111 (2013).
12. Cnudde, V., Boone, M. N. High-resolution X-ray computed tomography in geosciences: A review of the current technology and applications. *Earth-Science Reviews* **123**, 1-17 (2013).

13. Duchesne, M. J., Moore, F., Long, B. F., Labrie, J. A rapid method for converting medical Computed Tomography scanner topogram attenuation scale to Hounsfield Unit scale and to obtain relative density values. *Engineering Geology* **103**, 100-105 (2009).
14. Francus, P. *Image analysis, sediments and paleoenvironments*. (Springer Science & Business Media, 2006).
15. Croudace, I. W., Löwemark, L., Tjallingii, R., Zolitschka, B. Current perspectives on the capabilities of high resolution XRF core scanners. *Quaternary International* **514**, 5-15 (2019).
16. Croudace, I. W., Rindby, A., Rothwell, R. G. ITRAX: description and evaluation of a new multi-function X-ray core scanner. *Geological Society, London, Special Publications* **267**, 51-63 (2006).
17. Rothwell, R. G., Rack, F. R. New techniques in sediment core analysis: an introduction. *Geological Society, London, Special Publications* **267**, 1-29 (2006).
18. Zander, P. D., Żarczyński, M., Tylmann, W., Rainford, S.-k., Grosjean, M. Seasonal climate signals preserved in biochemical varves: insights from novel high-resolution sediment scanning techniques. *Climate of the Past* **17**, 2055-2071 (2021).
19. Zander, P. D., *et al.* A high-resolution record of Holocene primary productivity and water-column mixing from the varved sediments of Lake Żabińskie, Poland. *Science of the total environment* **755**, 143713 (2021).
20. Butz, C., Grosjean, M., Fischer, D., Wunderle, S., Tylmann, W., Rein, B. Hyperspectral imaging spectroscopy: a promising method for the biogeochemical analysis of lake sediments. *Journal of Applied Remote Sensing* **9**, 096031 (2015).
21. Müller, A. C., Guido, S. *Introduction to Machine Learning with Python: A Guide for Data Scientists*, 1 edn. (O'Reilly Media, 2016).
22. Jordan, M. I., Mitchell, T. M. Machine learning: Trends, perspectives, and prospects. *Science* **349**, 255-260 (2015).
23. Alpaydin, E. *Introduction to Machine Learning (Adaptive Computation and Machine Learning series)*, 3rd edn. (The MIT Press, 2014).
24. Hastie, T., Tibshirani, R., Friedman, J. *The elements of statistical learning: data mining, inference, and prediction*. (Springer Science & Business Media, 2009).
25. Kuwatani, T., *et al.* Machine-learning techniques for geochemical discrimination of 2011 Tohoku tsunami deposits. *Scientific Reports* **4**, 7044 (2014).
26. Murphy, K. P. *Machine learning: a probabilistic perspective*. (MIT press, 2012).

27. Cichocki, A., Phan, A.-H. Fast local algorithms for large scale nonnegative matrix and tensor factorizations. *IEICE transactions on fundamentals of electronics, communications and computer sciences* **92**, 708-721 (2009).
28. Févotte, C., Idier, J. Algorithms for nonnegative matrix factorization with the β -divergence. *Neural computation* **23**, 2421-2456 (2011).
29. Schubert, E., Sander, J., Ester, M., Kriegel, H. P., Xu, X. DBSCAN revisited, revisited: why and how you should (still) use DBSCAN. *ACM Transactions on Database Systems (TODS)* **42**, 1-21 (2017).
30. Bulian, F., *et al.* Multi-proxy reconstruction of Holocene paleoenvironments from a sediment core retrieved from the Wadden Sea near Norderney, East Frisia, Germany. *Estuarine, Coastal and Shelf Science* **225**, 106-251 (2019).
31. Gebhardt, A. C., *et al.* Petrophysical characterization of the lacustrine sediment succession drilled in Lake El'gygytgyn, Far East Russian Arctic. *Climate of the Past* **9**, 1933-1947 (2013).
32. Castelletti, A., Galelli, S., Restelli, M., Soncini-Sessa, R. Tree-based reinforcement learning for optimal water reservoir operation. *Water Resources Research* **46**, (2010).
33. Mnih, V., *et al.* Human-level control through deep reinforcement learning. *Nature* **518**, 529-533 (2015).
34. Sutton, R. S., Barto, A. G. *Reinforcement Learning: An introduction*. (MIT Press, 1998).
35. Ai, X., Wang, H., Sun, B. Automatic identification of sedimentary facies based on a support vector machine in the Arysium Graben, Kazakhstan. *Applied Sciences* **9**, 4489 (2019).
36. Bolandi, V., Kadkhodaie, A., Farzi, R. Analyzing organic richness of source rocks from well log data by using SVM and ANN classifiers: A case study from the Kazhdumi formation, the Persian Gulf basin, offshore Iran. *Journal of Petroleum Science and Engineering* **151**, 224-234 (2017).
37. Wrona, T., Pan, I., Gawthorpe, R. L., Fossen, H. Seismic facies analysis using machine learning. *Geophysics* **83**, 83-95 (2018).
38. Bolton, M. S. M., *et al.* Machine learning classifiers for attributing tephra to source volcanoes: an evaluation of methods for Alaska tephtras. *Journal of Quaternary Science* **35**, 81-92 (2020).
39. Jacq, K., *et al.* Sedimentary structures discriminations with hyperspectral imaging on sediment cores. *Computers and Geosciences*, (2020).

40. Ben Dor, Y., *et al.* Varves of the Dead Sea sedimentary record. *Quaternary Science Reviews* **215**, 173-184 (2019).
41. Behre, K. E. Coastal development, sea-level change and settlement history during the later Holocene in the Clay District of Lower Saxony (Niedersachsen), northern Germany. *Quaternary International* **112**, 37-53 (2004).
42. Davies, S. J., Lamb, H. F., Roberts, S. J. Micro-XRF Core Scanning in Palaeolimnology: Recent Developments. in *Micro-XRF Studies of Sediment Cores: Applications of a non-destructive tool for the environmental sciences* (ed. Croudace, I. W., Rothwell, R. G.) 189-226 (Springer Netherlands, 2015).
43. Rothwell, R. G., Croudace, I. W. Twenty years of XRF core scanning marine sediments: What do geochemical proxies tell us? in *Micro-XRF Studies of Sediment Cores: Applications of a non-destructive tool for the environmental sciences* (ed. Croudace, I. W., Rothwell, R. G.) 25-102 (Springer, 2015).
44. Anderton, R. Clastic facies models and facies analysis. *Geological Society, London, Special Publications* **18**, 31-47 (1985).
45. Daidu, F., Yuan, W., Min, L. Classifications, sedimentary features and facies associations of tidal flats. *Journal of Palaeogeography* **2**, 66-80 (2013).
46. Ltd, G. Geotek X-ray core imaging with CT. in (2018).
47. Hounsfield, G. N. Computerized transverse axial scanning (tomography): Part 1. Description of system. *The British journal of radiology* **46**, 1016-1022 (1973).
48. Zolitschka, B., Mingram, J., Gaast, S. V. D., Jansen, J., Naumann, R. Sediment logging techniques. in *Tracking environmental change using lake sediments* 137-153 (Springer, 2002).
49. Vos, P. C., Van Kesteren, W. P. The long-term evolution of intertidal mudflats in the northern Netherlands during the Holocene; natural and anthropogenic processes. *Continental Shelf Research* **20**, 1687-1710 (2000).
50. Meijles, E. W., *et al.* Holocene relative mean sea-level changes in the Wadden Sea area, northern Netherlands. *Journal of Quaternary Science* **33**, 905-923 (2018).
51. Dijkema, K. S., Reineck, H. E., Wolff, W. J., Group, W. S. W. *Geomorphology of the Wadden Sea Area: Final Report of the Section "Geomorphology" of the Wadden Sea Working Group.* (Balkema, 1980).
52. Ehlers, J. *Morphodynamics Wadden Sea*, 1 edn. (A.A. Balkema, 1988).

53. Dellwig, O., Hinrichs, J., Hild, A., Brumsack, H. J. Changing sedimentation in tidal flat sediments of the southern north sea from the holocene to the present: A geochemical approach. *Journal of Sea Research* **44**, 195-208 (2000).
54. Kolditz, K., Dellwig, O., Barkowski, J., Badewien, T. H., Freund, H., Brumsack, H. J. Geochemistry of salt marsh sediments deposited during simulated sea-level rise and consequences for recent and Holocene coastal development of NW Germany. *Geo-Marine Letters* **32**, 49-60 (2012).
55. Kolditz, K., *et al.* Geochemistry of Holocene salt marsh and tidal flat sediments on a barrier island in the southern North Sea (Langeoog, North-west Germany). *Sedimentology* **59**, 337-355 (2012).
56. Dellwig, O., *et al.* Geochemical and microfacies characterization of a Holocene depositional sequence in northwest Germany. *Organic Geochemistry* **29**, 1687-1699 (1998).
57. Dellwig, O., Watermann, F., Brumsack, H. J., Gerdes, G. High-resolution reconstruction of a holocene coastal sequence (NW Germany) using inorganic geochemical data and diatom inventories. *Estuarine, Coastal and Shelf Science* **48**, 617-633 (1999).
58. Freund, H., Gerdes, G., Streif, H., Dellwig, O., Watermann, F. The indicative meaning of diatoms, pollen and botanical macro fossils for the reconstruction of palaeoenvironments and sea-level fluctuations along the coast of Lower Saxony; Germany. *Quaternary International* **112**, 71-87 (2004).
59. Schüttenhelm, R. T. E., Laban, C. Heavy minerals, provenance and large scale dynamics of seabed sands in the Southern North Sea: Baak's (1936) heavy mineral study revisited. *Quaternary International* **133-134**, 179-193 (2005).
60. Beck, M., *et al.* Imprint of past and present environmental conditions on microbiology and biogeochemistry of coastal Quaternary sediments. *Biogeosciences* **8**, 55-68 (2011).
61. Dellwig, O., Böttcher, M. E., Lipinski, M., Brumsack, H. J. Trace metals in Holocene coastal peats and their relation to pyrite formation (NW Germany). *Chemical Geology* **182**, 423-442 (2002).
62. Flemming, B., Ziegler, K. High-resolution grain size distribution patterns and textural trends in the backbarrier environment of Spiekeroog Island (southern North Sea). *Senckenbergiana maritima* **26**, 1-24 (1995).
63. Flemming, B. W., Nyandwi, N. Land reclamation as a cause of fine-grained sediment depletion in backbarrier tidal flats (Southern North Sea). *Netherlands Journal of Aquatic Ecology* **28**, 299-307 (1994).

64. Hinrichs, J., Dellwig, O., Brumsack, H. J. Lead in sediments and suspended particulate matter of the German Bight: Natural versus anthropogenic origin. *Applied Geochemistry* **17**, 621-632 (2002).
65. Instruments, B. MS2/MS3 Magnetic Susceptibility System Manual. in) (1999).
66. de Montety, L., Long, B., Desrosiers, G., Crémer, J., Locat, J., Stora, G. Scanner use for sediment study: the influence of physical parameters, chemistry and biology on tomographic intensities. *Canadian Journal of Earth Sciences* **40**, 937-948 (2003).
67. Stalling, D., Westerhoff, M., Hege, H.-C. Amira: A highly interactive system for visual data analysis. *The visualization handbook* **38**, 749-767 (2005).
68. Aitchison, J. The statistical analysis of compositional data. *Journal of the Royal Statistical Society: Series B (Methodological)* **44**, 139-160 (1982).
69. Lee, A.-S., Huang, J.-J. S., Burr, G., Kao, L. C., Wei, K.-Y., Liou, S. Y. H. High resolution record of heavy metals from estuary sediments of Nankan River (Taiwan) assessed by rigorous multivariate statistical analysis. *Quaternary International* **527**, 44-51 (2019).
70. Weltje, G. J., Tjallingii, R. Calibration of XRF core scanners for quantitative geochemical logging of sediment cores: Theory and application. *Earth and Planetary Science Letters* **274**, 423-438 (2008).
71. Loska, K., Wiechuła, D. Application of principal component analysis for the estimation of source of heavy metal contamination in surface sediments from the Rybnik Reservoir. *Chemosphere* **51**, 723-733 (2003).
72. Schwestermann, T., *et al.* Multivariate Statistical and Multiproxy Constraints on Earthquake-Triggered Sediment Remobilization Processes in the Central Japan Trench. *Geochemistry, Geophysics, Geosystems* **21**, e2019GC008861 (2020).
73. Pedregosa, F., *et al.* Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **12**, 2825-2830 (2011).
74. Kaiser, H. F. The application of electronic computers to factor analysis. *Educational and psychological measurement* **20**, 141-151 (1960).
75. Nederbragt, A. J., Thurow, J. W. Digital sediment colour analysis as a method to obtain high resolution climate proxy records. in *Image analysis, sediments and paleoenvironments* 105-124 (Springer, 2005).
76. Berns, R. S. *Billmeyer and Saltzman's Principles of Color Technology*, 3rd edn. (Wiley-Interscience, 2000).

77. Van Der Walt, S., *et al.* Scikit-image: Image processing in python. *PeerJ* **2**, e453 (2014).
78. Silverman, B. W. *Density Estimation for Statistics and Data Analysis*. (Chapman & Hall Ltd., 1986).
79. Hunter, J. D. Matplotlib: A 2D graphics environment. *Computing in Science and Engineering* **9**, 99-104 (2007).
80. McKinney, W. Data Structures for Statistical Computing in Python. in *The 9th Python in Science Conference* 56-61 (2010).
81. Millman, K. J., Aivazis, M. Python for scientists and engineers. *Computing in Science & Engineering* **13**, 9-12 (2011).
82. Van Der Walt, S., Colbert, S. C., Varoquaux, G. The NumPy array: A structure for efficient numerical computation. *Computing in Science and Engineering* **13**, 22-30 (2011).
83. Virtanen, P., *et al.* SciPy 1.0--Fundamental Algorithms for Scientific Computing in Python. *Nature Methods* **17**, 261-272 (2020).
84. Capperucci, R. M., Bartholomä, A., Bungenstock, F., Enters, D., Karle, M., Wehrmann, A. The WASA core catalogue of Late Quaternary depositional sequences in the central Wadden Sea-a manual for the core repository. *Netherlands Journal of Geosciences*, (2022).
85. Wang, M., *et al.* A 600-year flood history in the Yangtze River drainage: comparison between a subaqueous delta and historical records. *Chinese Science Bulletin* **56**, 188-195 (2011).
86. Sluijs, A., *et al.* Arctic late Paleocene–early Eocene paleoenvironments with special emphasis on the Paleocene-Eocene thermal maximum (Lomonosov Ridge, Integrated Ocean Drilling Program Expedition 302). *Paleoceanography* **23**, (2008).
87. Keppler, F., Biester, H. Peatlands: A major sink of naturally formed organic chlorine. in *Chemosphere*. Elsevier Ltd (2003).
88. Leri, A. C., Myneni, S. C. B. Natural organobromine in terrestrial ecosystems. *Geochimica et Cosmochimica Acta* **77**, 1-10 (2012).
89. Zolitschka, B., *et al.* Pleistocene climatic and environmental variations inferred from a terrestrial sediment record—the Rodderberg Volcanic Complex near Bonn, Germany. *Zeitschrift der Deutschen Gesellschaft für Geowissenschaften* **165**, 407-424 (2014).

90. Jaccard, S., *et al.* Subarctic Pacific evidence for a glacial deepening of the oceanic respired carbon pool. *Earth and Planetary Science Letters* **277**, 156-165 (2009).
91. Ziegler, M., Lourens, L. J., Tuenter, E., Reichart, G.-J. Anomalously high Arabian Sea productivity conditions during MIS 13. *Climate of the Past Discussions* **5**, 1989-2018 (2009).
92. Schaumann, R. M., *et al.* The Middle Pleistocene to early Holocene subsurface geology of the Norderney tidal basin: new insights from core data and high-resolution sub-bottom profiling (Central Wadden Sea, southern North Sea). *Netherlands Journal of Geosciences* **100**, e15 (2021).
93. Dellwig, O., Watermann, F., Brumsack, H. J., Gerdes, G., Krumbein, W. E. Sulphur and iron geochemistry of Holocene coastal peats (NW Germany): A tool for palaeoenvironmental reconstruction. *Palaeogeography, Palaeoclimatology, Palaeoecology* **167**, 359-379 (2001).
94. Alve, E., Murray, J. W. Experiments to determine the origin and palaeoenvironmental significance of agglutinated foraminiferal assemblages. in *Proc. IV Inter. Workshop on Agglutinated Foraminifera. Krakow: Grzybowski Foundation* 1-11 (1995).
95. Nyquist, H. Certain Topics in Telegraph Transmission Theory. *Transactions of the American Institute of Electrical Engineers* **47**, 617-644 (1928).
96. Shannon, C. E. Communication in the Presence of Noise. *Proceedings of the IRE* **37**, 10-21 (1949).
97. Francus, P., Pirard, E. Testing for sources of errors in quantitative image analysis. in *Image analysis, sediments and paleoenvironments* 87-102 (Springer, 2005).
98. Charman, D. J. Peat and Peatlands. in *Encyclopedia of Inland Waters*. Elsevier Inc. (2009).
99. Mack, G. H., James, W. C., Monger, H. C. Classification of paleosols. *Geological Society of America Bulletin* **105**, 129-136 (1993).
100. Reinick, H. E., Wunderlich, F. Classification and origin of flaser and lenticular bedding. *Sedimentology* **11**, 99-104 (1968).
101. Karle, M., Bungenstock, F., Wehrmann, A. Holocene coastal landscape development in response to rising sea level in the Central Wadden Sea coastal region. *Netherlands Journal of Geosciences* **100**, e12 (2021).
102. Sheldon, N. D., Tabor, N. J. Quantitative paleoenvironmental and paleoclimatic reconstruction using paleosols. *Earth-Science Reviews* **95**, 1-52 (2009).

103. Wu, S., *et al.* Artificial intelligence reveals environmental constraints on colour diversity in insects. *Nature Communications* **10**, (2019).
104. Yu, Y., *et al.* Machine learning–based observation-constrained projections reveal elevated global socioeconomic risks from wildfire. *Nature communications* **13**, 1-11 (2022).
105. Basu, T., *et al.* Automated facies estimation from integration of core, petrophysical logs, and borehole images. in *AAPG Annual Meeting* 1-7 (2002).
106. Bittmann, F., Bungenstock, F., Wehrmann, A. Drowned palaeo-landscapes: archaeological and geoscientific research at the southern North Sea coast. *Netherlands Journal of Geosciences* **101**, e3 (2022).
107. Martin-Puertas, C., Tjallingii, R., Bloemsmas, M., Brauer, A. Varved sediment responses to early Holocene climate and environmental changes in Lake Meerfelder Maar (Germany) obtained from multivariate analyses of micro X-ray fluorescence core scanning data. *Journal of Quaternary Science* **32**, 427-436 (2017).
108. Lintern, A., *et al.* Sediment cores as archives of historical changes in floodplain lake hydrology. *Science of the Total Environment* **544**, 1008-1019 (2016).
109. Miller, H., Croudace, I. W., Bull, J. M., Cotterill, C. J., Dix, J. K., Taylor, R. N. A 500 year sediment lake record of anthropogenic and natural inputs to Windermere (English Lake District) using double-spike lead isotopes, radiochronology, and sediment microanalysis. *Environmental science & technology* **48**, 7254-7263 (2014).
110. Wilkinson, M. D., *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* **3**, 160018 (2016).
111. Bittmann, F., Bungenstock, F., Wehrmann, A. Drowned palaeo-landscapes: archaeological and geoscientific research at the southern North Sea coast. *Netherlands Journal of Geosciences* **101**, (2022).
112. Schlütz, F., Enters, D., Bittmann, F. From dust till drowned: the Holocene landscape development at Norderney, East Frisian Islands. *Netherlands Journal of Geosciences* **100**, e7 (2021).
113. Ng, A. *Machine learning yearning*. (deeplearning.ai, 2018).
114. Lee, A.-S., Enters, D., Titschack, J., Zolitschka, B. Facies characterisation of sediments from the East Frisian Wadden Sea (Germany): new insights from down-core scanning techniques. *Netherlands Journal of Geosciences* **100**, e8 (2021).

115. Weltje, G. J., Bloemsma, M. R., Tjallingii, R., Heslop, D., Röhl, U., Croudace, I. W. Prediction of Geochemical Composition from XRF Core Scanner Data: A New Multivariate Approach Including Automatic Selection of Calibration Samples and Quantification of Uncertainties. in *Micro-XRF Studies of Sediment Cores: Applications of a non-destructive tool for the environmental sciences* (ed. Croudace, I. W., Rothwell, R. G.) 507-534 (Springer Netherlands, 2015).
116. Pawlowsky-Glahn, V., Egozcue, J. J. Compositional data and their analysis: An introduction. *Geological Society Special Publication* **264**, 1-10 (2006).
117. Abu-Mostafa, Y. S., Magdon-Ismail, M., Lin, H.-T. *Learning From Data: A Short Course*. (AMLBook, 2012).
118. Morales, J. L., Nocedal, J. Remark on “algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound constrained optimization”. *ACM Trans Math Softw* **38**, Article 7 (2011).
119. Breiman, L. Random forests. in *Machine Learning* 5-32 (Springer, 2001).
120. Chang, C.-C., Lin, C.-J. LIBSVM: A Library for Support Vector Machines. *ACM Transactions on Intelligent Systems and Technology* **2**, 1-27 (2011).
121. King, G., Zeng, L. Logistic Regression in Rare Events Data. *Political Analysis* **9**, 137-163 (2001).
122. Brodersen, K. H., Ong, C. S., Stephan, K. E., Buhmann, J. M. The balanced accuracy and its posterior distribution. in *International Conference on Pattern Recognition* 3121-3124 (2010).
123. Kelleher, J. D., Mac Namee, B., D'arcy, A. *Fundamentals of machine learning for predictive data analytics: algorithms, worked examples, and case studies*. (MIT press, 2020).
124. Harris, C. R., *et al.* Array programming with NumPy. *Nature* **585**, 357-362 (2020).
125. Millman, K. J., Aivazis, M. Python for scientists and engineers. *Computing in Science and Engineering* **13**, 9-12 (2011).
126. Tjallingii, R., Röhl, U., Kölling, M., Bickert, T. Influence of the water content on X-ray fluorescence core-scanning measurements in soft marine sediments. *Geochemistry, Geophysics, Geosystems* **8**, (2007).
127. Böning, P., Bard, E., Rose, J. Toward direct, micron-scale XRF elemental maps and quantitative profiles of wet marine sediments. *Geochemistry, Geophysics, Geosystems* **8**, (2007).

128. Jansen, J., Van der Gaast, S., Koster, B., Vaars, A. CORTEX, a shipboard XRF-scanner for element analyses in split sediment cores. *Marine geology* **151**, 143-153 (1998).
129. Max, L., *et al.* Sea surface temperature variability and sea-ice extent in the subarctic northwest Pacific during the past 15,000 years. *Paleoceanography* **27**, n/a-n/a (2012).
130. Benz, V. a. E. O. a. G. R. a. L. F. a. T. R. Last Glacial Maximum sea surface temperature and sea-ice extent in the Pacific sector of the Southern Ocean. *Quaternary Science Reviews* **146**, 216--237 (2016).
131. Lamy, F., *et al.* Increased Dust Deposition in the Pacific Southern Ocean During Glacial Periods. *Science* **343**, 403-407 (2014).
132. Ullermann, J., Lamy, F., Ninnemann, U., Lembke-Jene, L., Gersonde, R., Tiedemann, R. Pacific-Atlantic Circumpolar Deep Water coupling during the last 500 ka. *Paleoceanography* **31**, 639-650 (2016).
133. Wang, W., *et al.* Dating North Pacific Abyssal Sediments by Geomagnetic Paleointensity: Implications of Magnetization Carriers, Plio-Pleistocene Climate Change, and Benthic Redox Conditions. *Frontiers in Earth Science*, 577 (2021).
134. Nürnberg, D. RV SONNE Fahrtbericht / Cruise Report SO264 - SONNE-EMPEROR: The Plio/Pleistocene to Holocene development of the pelagic North Pacific from surface to depth – assessing its role for the global carbon budget and Earth's climate, Suva (Fiji) – Yokohama (Japan), 30.6. – 24.8.2018. in *GEOMAR Report*). GEOMAR Helmholtz-Zentrum für Ozeanforschung (2018).
135. Chao, W.-S., *et al.* Glacial-interglacial variations in productivity and carbonate deposition in the Northwest Pacific during the last 500,000 years. *Frontiers in Earth Science*, (In prep.).
136. Clifford, G. Chapter 15-BLIND SOURCE SEPARATION: Principal & Independent Component Analysis. in *Biomedical Signal and Image Processing*). MIT Electrical Engineering and Computer Science (2008).
137. Geurts, P., Ernst, D., Wehenkel, L. Extremely randomized trees. *Machine learning* **63**, 3-42 (2006).