# Exploring Spectral Imaging as a Tool for Stratigraphic Analysis

Dissertation
to obtain the academic degree
Doctor rerum naturalium

submitted to the
Department of Earth Sciences
Freie Universität Berlin

by

## Johann Joseph Vincent Haburaj

Berlin 2020

# Acknowledgements

# Abstract

This thesis examines how quantitative colour measurements and spectral recordings can be utilised to add traceability to the process of stratigraphic analysis carried out for archaeological and geoscientific research. The non-destructive, quantitative and often extensive character of this data holds much potential for stratigraphic documentation and interpretation. Quantitative methods are examined, which complement traditional perception-based documentation and interpretation of stratigraphies without the need of complex laboratory work. Likewise, it is shown how the results of such laboratory work can be further utilised by combining them with spectral imaging techniques.

Digital photography, as carried out as a standard procedure during fieldwork, constitutes a basic form of spectral imaging. Throughout this thesis, digital photographs systematically recorded with a common RGB camera are therefore treated as physical measurements. By analysing RGB, multispectral and hyperspectral image data of soil and sediment profiles, it is shown, that digital image data offers a reliable basis for the transparent and reproducible delineation of stratigraphic layers. The produced stratigraphic sequences based on spectral data are of similar quality as sequences based on traditional sedimentological data obtained from laboratory work. Spectral data, however, can be acquired during fieldwork and offers a fast and non-destructive alternative to exhaustive laboratory analyses.

Furthermore, it is shown that the extensive character of spectral image data can be utilised to produce maps of chemical and physical parameters along soil and sediment profiles. Through the application of statistical models which involve the spectral image data and sediment properties obtained from samples analysed in the laboratory, reliable maps are created for various parameters like soil organic matter or iron oxide contents. The presented results show that RGB images as well as hyperspectral images can be utilised for this purpose.

Exploring possible applications of spectral imaging, the results presented in this thesis promote a more transparent and reproducible documentation for often destructive archaeological and geoscientific fieldwork.

# Zusammenfassung

In der vorliegenden Arbeit wir untersucht, wie quantitative Farbmessungen und Spektralaufnahmen dazu beitragen können, den Prozess der stratigraphischen Analyse im Kontext archäologischer und geowissenschaftlicher Forschung transparenter und objektiver zu gestalten. Solche Daten bieten durch ihren zerstörungsfreien, quantitativen und oftmals flächenhaften Charakter eine vielversprechende Informationsquelle für die stratigraphische Dokumentation und Interpretation. In dieser Arbeit werden daher quantitative Methoden untersucht, welche die traditionelle, auf Wahrnehmung basierende Dokumentation und Interpretation von Stratigraphien ergänzen können, ohne dabei komplexe Laborarbeit zu erfordern. Weiterhin wird gezeigt, wie die Ergebnisse herkömmlicher Labormessungen mit bildgebenden Spektralmessungen kombiniert werden können um eine belastbare Datengrundlage für stratigraphische Analysen zu schaffen.

Digitale Fotografie ist heutzutage ein fester Bestandteil archäologischer und geowissenschaftlicher Feldarbeit. Gleichzeitig stellt eine digitale RGB Fotografie eine einfache Art von Spektralmessung dar. In der vorliegenden Arbeit werden die systematisch durchgeführten Aufnahmen einer herkömmlichen RGB Kamera daher als physikalische Messungen behandelt. Es wird gezeigt, dass die Analyse solch digitaler Bilddaten eine verlässliche Grundlage für die transparente und reproduzierbare Abgrenzung von stratigraphischen Schichten bildet. Hierbei wird neben RGB Daten auch der Mehrwert von multi- und hyperspektralen Bilddaten untersucht. Die so produzierten stratigraphischen Sequenzen sind von ähnlicher Qualität wie Sequenzen, die auf herkömmlichen sedimentologischen Laborarbeiten basieren. Spektrale Messungen können im Vergleich zu vielen anderen sedimentologischen Messungen jedoch häufig bereits während der Feldarbeiten aufgenommen werden und bieten daher eine schnelle und zerstörungsfreie Alternative zu aufwändigen Laborarbeiten.

Weiterhin wird in dieser Arbeit gezeigt, dass spektrale Bilddaten dazu genutzt werden können, chemische und physikalische Eigenschaften entlang von Boden- und Sedimentprofilen zu kartieren. Zuverlässige Kartierungen zahlreicher Parameter wie beispielsweise dem Organik- oder Eisenoxidgehalt konnten durch statistische Modellierungen erzeugt werden. Sowohl RGB als auch hyperspektrale Bilddaten können für diesen Zweck verwendet werden. Die so erzeugten Kartierungen bilden eine verlässliche Datengrundlage für die Dokumentation und Interpretation stratigraphischer Sequenzen.

Die Ergebnisse dieser Arbeit fördern eine transparente und reproduzierbare Dokumentationsweise, welche insbesondere im Kontext oftmals zerstörerischer archäologischer und geowissenschaftlicher Feldarbeit an Bedeutung gewinnt.

# Contents

# List of Figures

# List of Tables

# CHAPTER 1

## Introduction

Stratigraphic documentation and interpretation are crucial parts of archaeological and geoscientific research. While the topical focus of the two disciplines often differs, major fieldwork methods are very similar: fieldwork standards like drawing and measuring, which were established in the 20th century, today are complemented by a growing number of quantitative and digital methods. Nevertheless, the human reading of geoscientific and archaeological sediment layers and sections under fieldwork conditions relies heavily on professional expertise (including both technical and theoretical knowledge) and visual and tactile perception. These essential factors, however, differ significantly between individual researchers. The narrative of a stratigraphy is therefore strongly influenced by subjective factors in the course of its excavation and documentation. Following a recent demand for more transparent and reproducible research approaches (e.g. Mesirov, 2010; Peng, 2011; Marwick, 2017; Marwick et al., 2017), this thesis examines how these drawbacks of the common methodological approach for stratigraphic analysis can be overcome by including an additional layer of quantitative data into the process of documentation. More specifically, the presented studies examine how systematically recorded digital image data as well as extensive and selective spectral recordings could extend the data basis for stratigraphic documentation and interpretation.

As systematically recorded digital images capture the interaction of light with the depicted objects, they are treated as quantitative physical measurements throughout this thesis. Such extensive optical measurements (RGB, multispectral or hyperspectral image data) can be used to capture whole profiles or plana during fieldwork. As shown in this thesis, the quantitative character of this data allows for statistical analyses which complement the traditional methodology by adding traceability, transparency and objectivity to the stratigraphic record. Moreover, the presented studies indicate, that even selectively measured spectral data (point spectrometer) can be used for this purpose. Both extensive and selective optical data are easy and fast to acquire during fieldwork and offer an inexpensive alternative to exhaustive laboratory analyses. Whilst the focus of this thesis lies on archaeological fieldwork, many of the findings are equally applicable in a geoscientific context.

The studies included in this thesis were carried out in the context of Research Area A *Spatial Environment and Conceptual Design* and Research Area B *Monumentalized Knowledge. Extra-Large Projects in Ancient Civilizations* of the *Cluster of Excellence EXC264 Topoi*. Collaboration between researchers of the two research areas allowed for a variety of study locations which all comprise an existent and profound stratigraphic record. These records are based on long-term fieldwork, expert knowledge and often time- and cost-intensive quantitative laboratory analyses. Hence, they provided a reliable basis from which the three presented studies could asses the applicability of the proposed methodology as well as the additional value provided by the systematic analysis of digital image data and spectral recordings acquired during fieldwork.

## 1.1   Objectives

Building upon recent technological advancements in the fields of spectral imaging and image analysis, this thesis explores the question, whether spectral recordings (RGB, multispectral and hyperspectral data) can be utilised as a reliable basis for stratigraphic documentation and interpretation. The presented investigations target to examine the questions, (i) whether classification algorithms can be utilised to delineate stratigraphic layers based on RGB, multispectral and hyperspectral data recorded during fieldwork and in the laboratory, (ii) how spectral recordings perform when compared to often time-consuming and expensive laboratory analyses, and (iii) if spectral image data recorded during fieldwork can be used to produce reliable maps of sediment properties along archaeological profiles. These aspects are examined in the context of ongoing archaeological and geoscientific projects with an existent and profound stratigraphic record, which is used for controlling the quality of the respective results. It is shown how the proposed methodology can be used to reproduce, complement and extend the existent stratigraphic record. The presented results indicate that optical data, as used throughout the presented studies, should be regarded a valuable method for stratigraphic documentation and interpretation: it can serve as a reference for the visual perception of the researcher at work (chapter 4), can offer an alternative to exhaustive laboratory analyses in specific cases (chapter 5) or can be used to further exploit the results of such laboratory analyses (chapter 6).

Past studies already showed promising results of spectroscopic measurements as a tool for stratigraphic documentation of soil profiles (e.g. Steffens and Buddenbaum, 2013; Hobley et al., 2018). As these studies are, however, limited to laboratory hyperspectral imaging, a major research goal of this thesis is the evalutaion of the on-site applicability of spectral imaging during archaeological and geoscientific fieldwork. The potential of image analysis techniques and high-quality hyperspectral data for fieldwork is investigated.

One major topic of this thesis is the delineation of stratigraphic layers based on the statistical analysis of image data and spectral recordings and is published within two research papers:

- **Evaluating the Potential of Semi-Automated Image Analysis for Delimiting Soil and Sediment Layers** by V. Haburaj, Krause J., Pless S., Waske B. & Schütt B. (2019, Journal of Field Archaeology 44(8), 538-549; own contribution: 80 %).

  In order to assess the potential of semi-automated image classification as a tool for delineating

stratigraphic layers, this study compares the quality of layer delineations obtained from unsupervised classification of different image datasets acquired during fieldwork. High-resolution RGB images and hyperspectral images of lower spatial resolution were recorded. Benefits and drawbacks of the different spatial and spectral resolution as well as multiple pre-processing steps were examined. The stratigraphic layers delineated by the applied unsupervised image classifications were compared to the stratigraphic record based on expert knowledge and sedimentological data. The results show that the identification of stratigraphic features can be achieved using easy-to-apply RGB imaging.

- **On-Site VIS-NIR Spectral Reflectance and Colour Measurements – A Fast and Inexpensive Alternative for Delineating Sediment Layers Quantitatively? A Case Study from a Monumental Bronze Age Burial Mound (Seddin, Germany)** by V. Haburaj, Nykamp M., May J., Hoelzmann P. & Schütt B. (2020, Heritage 3(2), 528–548; own contribution: 75 %). Building upon Haburaj et al. (2019), this study explores if the proposed image analysis workflow can be transferred to more complex archaeological profiles and improved by further pre-processing of the data. Additionally, the chemical and physical properties of 31 sediment samples obtained from an archaeological profile were analysed in the laboratory. The sedimentological record obtained from this selectively measured data was used as input for statistical cluster analyses. Clustering results of (i) the measured element concentrations (ii) the measured visible and near-infrared spectral data of the samples and (iii) the values of image pixels covering the sampling areas were compared to assess the performance of spectroscopic measurements in relation to traditional laboratory-based techniques for sedimentological analysis. The outcome of this study shows that colour measurements and high-quality spectral data in general could offer a fast and inexpensive quantification of differences between individual stratigraphic layers. Furthermore, the acquired spectral image data helped to delineate and discuss the stratigraphic layers of the examined archaeological profile.

Systematic colour measurements and spectral measurements constitute quantitative physical data that is highly influenced by the chemical and physical composition of the examined object (e.g. Baumgardner et al., 1985). This connection between the properties of a material and its spectral reflectance is often used to create maps of specific properties based on spectral images, which then are, for example, used to assess the quality of soils or to support stratigraphic analysis. In the fields of archaeology and the geosciences this methodology is yet limited to laboratory studies (e.g. Steffens and Buddenbaum, 2012; Heil et al., 2020) and remote sensing (e.g. Retzlaff, 2015) and only recently was complemented by single on-site studies that feature exhaustive sampling strategies (Zhang and Hartemink, 2019a). Building upon these studies, the mapping of sediment properties based on RGB and hyperspectral images of complex archaeological profiles, acquired during fieldwork and paired with the analysis of a regular amount of sediment samples, was examined within a research article:

- **Coupling Spectral Imaging and Laboratory Analyses to Digitally Map Sediment Parameters and Stratigraphic Layers in Yeha, Ethiopia** by V. Haburaj, Japp S., Gerlach I., Hoelzmann P. & Schütt B. (2020, PLoS ONE 15(9): e0238894; own contribution: 85 %). This study examines if a workflow to extrapolate sediment properties spatially based on image

data can be transferred to complex archaeological profiles which are recorded under fieldwork conditions, as opposed to laboratory studies. A total of 16 sediment samples was analysed in the laboratory and used to train regression models involving RGB image data, RGB derived CIELAB (Commission Internationale de l'Éclairage) image data and hyperspectral image data. The image data was acquired during fieldwork and depicts two archaeological profiles. The obtained regression models were used to successfully map multiple sediment properties. By including two individual profiles and statistical cross-validation the accuracy and reproducibility of the results were verified.

The findings of the listed publications show clear benefits and shortcomings of the proposed methodology. Conclusions, recommendations for action and thoughts on future applications are presented in chapter 7.

## 1.2   Towards a gain of knowledge

Our criteria for knowledge are standards and rules that are accepted (or neglected) as *conventions* by purpose (Spinner, 1974, 11). Referring to Feyerabend, Spinner argues in his book *Pluralismus als Erkenntnismodell* (1974) that basic epistomological problems are not solved by proof but by *feasible* decisions. Regarding technological advances, this likewise means that new technologies can lead to a gain of knowledge. At the same time, Spinner (1974) sees the epistomological problem as a highly theoretical set of issues that *a priori* excludes a single definite solution. He therefore concludes that *pluralism* is always possible and necessary. In the same way, disciplinary borders should always be questioned (e.g. Ropohl, 2009, 2012). This is especially true regarding the archaeological discipline: by tradition it is strongly connected to other disciplines like ethnology, geology, chemistry, geography or linguistics. Archaeology, similar to many other disciplines, should therefore be understood as a discipline that involves several different disciplines, meaning it is *transdisciplinary* by nature. However, this does not mean that archaeology can not be considered as a discipline itself. Rather, we should not take *disciplines* as natural entities but as historically grown constructions (Mittelstraß, 2003; Meier and Tillessen, 2011). Meier and Tillessen (2011) for example observe that interdisciplinary archaeological projects (and most archaeological projects involve multiple disciplines) often discourse on the same problems, originating from different methodological competences and theoretical concepts that encounter each other. Their proposed interdisciplinary hermeneutic therefore incorporates, among others, a constant exchange between the disciplines on all levels. The contents of the thesis at hand should be considered as an approach to adapt this concept to the act of stratigraphic analysis in a pluralist way (Figure 1.1).

The critical role that pluralism ascribes to the *feasible* decisions can possibly be described as an replacement of the dichotomy truth/lie with feasible/impossible, as for example proposed by Piglia (1992) (or as Spinner (1974) on p.13 puts it (translated from german): "the omnipresent possibility of alternatives"). Scientific work is always determined by its contemporary technological capabilities. In the archaeological discipline as an intersection of humanities and sciences, we can observe a significant gain of knowledge through technological advancements in the last decades. Likewise this
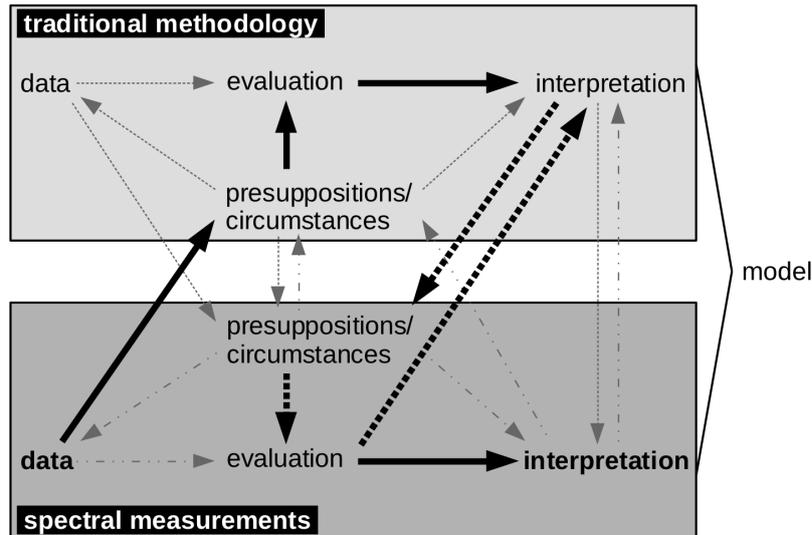
**Figure 1.1:** Integration of spectral measurements into the process of traditional stratigraphic documentation and interpretation based on an interdisciplinary research model (adapted from Meier and Tillessen 2011). While a constant exchange between the two entities is illustrated by grey lines, thick black lines illustrate the potential gain of knowledge. An exchange between (i) archaeological interpretation and (ii) presuppositions/circumstances and evaluation of spectral data takes place iteratively (thick dotted lines), leading to a final interpretation.

gain has to be seen critical. Harris (1979) already recognised, that the basis of archaeological work is a combination of practical expertise and theoretical background. Considering theoretical pluralism and fallibilism – as well as Feyerabend's *Against Method* (1975) – neither theory nor practical expertise or technological advances should be accepted as panacea for archaeological research. The sole solution remaining therefore is, to constantly test existing interpretations, opinions and hypotheses and by that move towards a possible gain of knowledge.

But how can this concept be transferred to stratigraphic interpretation? First, the analysis of digital image data and spectral data should be seen as a tool, emerging from technological advance and involving a possible gain of knowledge. Second, existing *conventions* should not be discarded. We rather suggest the proposed methodology as *possibility* to test the existing way of delineating diagnostic layers, which is based on time-consuming cognitive – and thus subjective-irrational – individual work. Third, a critical-rational attitude in combination with the proposed methodology allows for an increased certainty when delineating diagnostic layers and likewise holds potential information that may prove usable for the analysis of physical and chemical properties of these layers.

When Einstein (1955) discusses the evaluation of scientific theories, he distinguishes between *interior* (inner simplicity, consistency, systematic coherence) and *exterior* (confrontation with the real world) criticism (cf. Spinner, 1974). The complementary application of spectral imaging offers critique of the traditional stratigraphic interpretation from both within and outside: a confrontation of the cognitive interpretation with reality (actual spectral reflectance) and a critical evaluation of traditional conventions. Likewise the transdisciplinary character of stratigraphic interpretation is expanded (Ropohl, 2012).

Prior to a gain of knowledge, however, a critical analysis of the applicability and validity of *possible* methods has to be carried out. In this context, this thesis examines the applicability of spectral

imaging as a method of documentation and interpretation during archaeological and geoscientific fieldwork. The implementation into the traditional methodological conventions is carried out in the sense of Spinner's pluralism.

# CHAPTER 2

## State of the art

Documentation standards applied during archaeological and geoscientific excavations, which were established in the 20$^{th}$ century, are nowadays complemented by a growing number of on-site measurements and analyses carried out in various laboratories in the aftermath of an excavation. While this added layer of information leads to a more precise and transparent documentation of stratigraphic horizons, the on-site description and interpretation of the stratigraphic material is still influenced by the subjective perception of the respective researcher at work: the initial reading of the geoscientific and archaeological sediments under fieldwork conditions relies heavily on the researcher's visual and tactile perception. Additionally the basic documentation techniques used throughout this process are still based on non-reproducible methods like the manual characterization of the soil texture, technical drawings of sections and plana, and the numbering of layers and interfaces. Whilst these methods are widely accepted and have proven to produce a semi-uniform style of on-site documentation of archaeological and geoscientific excavations, the subjective and often non-reproducible character of these techniques indicates that assistive techniques are necessary to add transparency and reproducibility to the process of stratigraphic documentation and interpretation. To stress the necessity of the mentioned development, the following pages give an overview of the current state of stratigraphic documentation and interpretation as well as an introduction to the topic of spectroscopy. Historically, the two disciplines of archaeology and geoscience are heavily entangled: although their topical focus mostly differs, their research objects and methodology are often similar or even the same. As this thesis mainly focuses on archaeological fieldwork methodology, the following overview primarily addresses standards and developments from the archaeologies.

## 2.1 The echo of the Harris Matrix

General principles of geology had an important impact on the development of archaeology especially from the nineteenth century onwards: concepts of stratification were adapted and used as a basis for the documentation and interpretation of archaeological stratigraphies (e.g. Daniel, 1976). In the early nineteen-sixties Wheeler (1954) and Kenyon (1952) proposed a systematic numbering of stratigraphic layers and also put emphasis on distinct interfaces. Nevertheless, until Harris (1979) the only stratigraphical rule that was recognised in archaeology was that the age of strata successively grows from top layers to bottom layers. This rule became later known as the *Law of Superposition.*

During the nineteenth century the most common technique for recording archaeological excavations was the preparation of a schematic diagram after an excavation. Harris (1979) calls the manual drawing of a section by Willett (1880) an "exception". By the middle of the twentieth century documentation of excavations had developed further, now including plans and drawings that mainly depict anthropogenic features as walls, pits and ditches. Soil layers were mainly neglected (Harris, 1979, 25-28). The first systematic approach for a more holistic recording of archaeological stratigraphies was introduced by Harris (1979). He proposed the now widely used *Harris Matrix* as an instrument for capturing relations between stratigraphic layers. In the form of a grid of rectangular boxes, connected by lines, it represents the stratigraphic sequence. The Harris Matrix is based on the following rules (or laws), defined by Harris (1979):

1. *The Law of Superposition*: "In a series of layers and interfacial features, as originally created, the upper units of stratification are younger and the lower are older, for each must have been deposited on, or created by the removal of, a pre-existing mass of archaeological stratification." (Harris, 1989, 30)

2. *The Law of Original Horizontality*: "Any archaeological layer deposited in an unconsolidated form will tend towards a horizontal position. Strata which are found with tilted surfaces were originally deposited that way, or lie in conformity with the contours of a pre-existing basin of deposition." (Harris, 1989, 31)

3. *The Law of Original Continuity*: "Any archaeological deposit, as originally laid down, or any interfacial feature, as originally created, will be bounded by a basin of deposition, or may thin down to a feather-edge. Therefore, if any edge of a deposit or interfacial feature is exposed in a vertical view, a part of its original extent must have been removed by excavation or erosion, and its continuity must be sought, or its absence explained." (Harris, 1989, 32)

4. *The Law of Stratigraphical Succession*: "A unit of archaeological stratification takes its place in the stratigraphic sequence of a site from its position between the undermost (or earliest) of the units which lie above it and the uppermost (or latest) of all the units which lie below it and with which the unit has a physical contact, all others superpositional relationships being redundant." (Harris, 1989, 34)

The sole goal of Harris' approach is the creation of a sequence and thus he is criticised often (Renfrew and Bahn, 1996; Warburton, 2003). Warburton (2003), for example, argues that Harris

was unable to accept the concept of 'living floors' (p.13). Furthermore a big disadvantage of Harris' methodology is the neglecting of sections, which can prove crucial when architecture or similar features cut through stratigraphic sequences. On page 24 Warburton (2003) emphasises that the sole establishment of a sequence is pointless if its reliability and purpose are not examined: one should also focus on the classification of the deposits when interpreting stratigraphy. Warburton's thoughts are hereby characteristic for a general development in archaeology: with better techniques of chemical and physical analysis the characterization of layers gains importance:

> "Defining the vertical and horizontal spatial distribution of stratigraphic and excavation units is the first step in conducting an excavation. Defining the nature of those deposits is a far more complicated process, but decisive for the excavation, as the sequence of the deposits does not alone allow conclusions about the material to be drawn." (Warburton, 2003, 57)

While the characterisation of layers may indeed be crucial for archaeological work, we often neglect the fact that *identifying* the layers remains highly subjective in many cases. The act of layer identification is mentioned but not further described in most reference books (e.g. Roskams, 2001). The methodology proposed by Harris and his successors proves right, useful and usable but likewise we should not stop to refine, question and enhance it.

## 2.2   The current state of stratigraphic analysis

The traditional way of identifying diagnostic soil horizons during archaeological excavations is still based on what Wheeler (1960) recommended: the moistening of soil profiles and careful scratching accompanied by a contemplation during different times of the day (due to changing lighting conditions). In his book *Principles of Archaeological Stratigraphy* Harris (1989) states that "an excavator must have a theory of archaeological stratigraphy in order to know what to observe and record on an archaeological excavation" (p.40). He refers to Pyddoke (1961) who argues that interpretation is something that "must be learned on excavations and not from handbooks". Harris points out that a good combination of practical experience and a theoretical knowledge of stratigraphic principles is necessary. He mentions the principle of *uniformitarianism* (ger. *Aktualismus*, see e.g. von Hoff, 1822) as an important factor when identifying stratigraphic processes and its components. Harris hereby shows that a trained eye in combination with a theoretical background is the only solution to stratigraphic interpretation. Likewise his thoughts show how stratigraphic analysis significantly differs from an objective method: personal practical experience varies enormously between researchers and the theoretical background only gives us a framework similar to a recommendation for action.

Besides Harris' work, several assistive technologies for stratigraphic interpretation are available nowadays. Many of these methods (e.g. chemical analyses of soil samples, absolute dating methods) require chemical or physical processing of soil and sediment samples and therefore should be considered destructive methods. In contrast, non-destructive techniques like the 3D documentation of

archaeological trenches (laserscanning, structure from motion) certainly hold potential for capturing stratigraphies in an extensive three-dimensional way, but likewise are only capable of capturing what already has been defined as a crisp border between horizons, as each layer is documented only after it is excavated. Regarding soil and sediment colour, the widely used system proposed by Munsell (1929) offers a systematic approach: through the coding of colour hues and shades a certain degree of objectivity is achieved. Along with digital photography (photogrammetry), technical drawings and an on-site characterisation of the soil or sediment texture (loosely based on the USDA soil texture triangle) the Munsell colour system, today, constitutes the basis of on-site stratigraphic analysis during many archaeological excavations.

In the aftermath of an excavation, the mentioned techniques are often complemented by quantitative analyses based on sampled soil or sediment. Chemical and physical properties of the materials (e.g. Hassan, 1978), the analysis of thin sections (e.g. French, 2003) and phytoliths (e.g. Meister et al., 2017), as well as absolute dating methods offer a large amount of data which can be used to characterise individual stratigraphic layers and discuss the relation between those layers. Together with the expert's knowledge regarding the study site, the temporal context of the stratigraphic layers and the on-site documentation, a final stratigraphic interpretation is produced.

While this procedure leads to a robust stratigraphic interpretation in most cases, there are two factors which imply difficulties: (i) the aforementioned subjectivity, which is involved in the original fieldwork, influences on-site documentation and sampling and therefore has a significant impact on the final interpretation, and (ii) the quantitative data used to discuss and interpret the stratigraphic record remains limited to the sampled materials and therefore has to be seen as selective data. In this context, digital imaging, which has become a standard procedure for documenting sections and plana, holds the potential to be analysed quantitatively and thereby support the identification and characterisation of layers during fieldwork and in the aftermath of an excavation. The extensive character of such digital image data and the treatment of this image data as physical measurements can help to compensate for the mentioned issues. The statistical analysis of image data (or optical data in general) may help researchers to overcome the subjectivity which is currently part of stratigraphic analysis.

## 2.3   Spectroscopy and spectral imaging

The main principle of spectroscopy as a method is the interaction of electromagnetic radiation with different objects. Electromagnetic radiation emitted by, for example, the sun or artificial light sources, is reflected by each material in a unique manner. Depending on the chemical and physical composition of its surface and inner material, each object reflects the incoming radiation (or light) differently. The reflection of an ideal perfect reflector would only be specular (Figure 2.1), whereas most real objects reflect a combination of specular and diffuse reflection. Figure 2.1 shows a simplified model of the interaction of incident light (radiation) with an object. In reality, the interaction is more complex than depicted in Figure 2.1, since a part of the incoming radiation penetrates the surface of the object and is reflected chaotically by the inner material of the object. Additionally, each object with
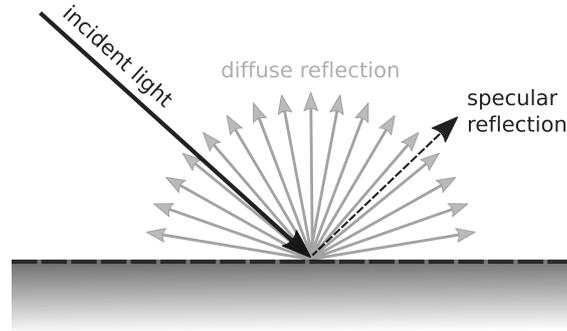
**Figure 2.1:** Spectral reflectance of an ideal plain surface. Measured reflectance is always a mixture of diffuse and specular reflection.

a temperature above absolute zero (0 K or $-274°$C) emits radiation on its own. The actual reflection, which is measured by a spectral sensor therefore always is a composite, resulting from multiple processes (cf. Miller, 2001). The reflected radiation most commonly is measured in % reflectance (R, where 100 % corresponds to an ideal perfect reflector) and is then often transformed to apparent absorbance $A = log(1/R)$ for further processing (Stenberg et al., 2010).

Spectral sensors, being optical sensors, only cover the part of the electromagnetic spectrum between a wavelength of approximately 200 nm and 20,000 nm. This covers the ultraviolet range (UV, 200 - 420 nm), the visible range (VIS, 420 - 780 nm), the near infrared range (NIR, 780 - 2,500 nm) and the infrared range (IR, 2,500 - 25,000 nm - 1 mm) (cf. Skrabal, 2009). Depending on the technology used, spectral imaging sensors capture different parts of these ranges with a different spectral and spatial resolution: while digital RGB imaging, similar to the human eye, only covers specific parts of the VIS part of the spectrum, most studies regarding soil and sediment analysis or remote sensing include a variable number of image bands from the VIS and NIR range. Depending on the number of bands such sensors are specified as multi- or hyperspectral sensors. While there is no uniform definition, that distinguishes multi- from hyperspectral sensors, the terms are roughly related to the actual number of acquired image bands. In the study at hand, the term multispectral is used to describe sensors which cover specific wavelengths and the term hyperspectral is used for sensors covering a specific spectral range with equidistant image bands.

The spectral imaging devices used throughout the presented studies capture data as raster stacks of the dimension *N* x *M* x *K*, where *N* and *M* are the spatial resolution of each acquired raster and *K* equals the number of spectral bands captured. Due to the three dimensions of the data, such raster stacks are often addressed as "cubes". A six-channel multispectral camera, for example, may capture a raster stack (or cube) of the following dimension: 1920 x 1200 x 6. Each band of this raster stack equals a greyscale image (1920 x 1200 pixels), which, for processing purposes, can be considered as a matrix – each pixel of the raster stack as a vector of length 6 which equals the measured spectral reflectance of the pixel (Figure 2.2c). The acquired data can therefore be processed using matrix or vector data analysis. The measured spectral signal often is used for image classification or the assessment of specific chemical or physical properties.

While spectral sensors often weigh several kilograms and require an external power supply, ongoing technological progress leads to smaller and more affordable devices which are more easy to use under fieldwork conditions. The presented studies in chapters 4, 5, and 6 examine the performance
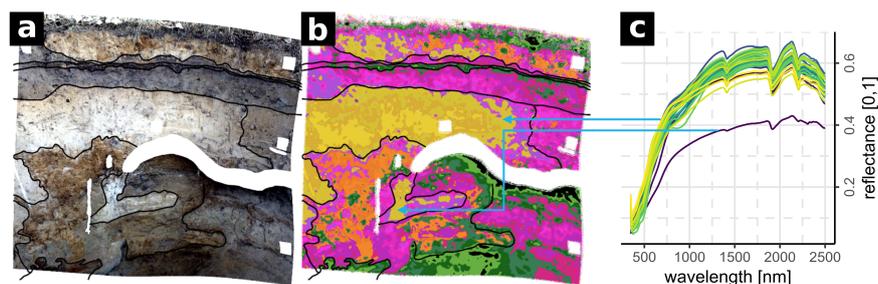
**Figure 2.2:** Conceptual framework: an archaeological or geoscientific profile (a) is captured with a hyperspectral camera, which allows one to delineate horizons based on quantitative image analyses (b) and assess multiple physical and chemical properties of the material based on the reflectance spectrum captured for each pixel (c).

of state-of-the-art snapshot devices (covering the VIS and NIR range) when they are used for delineating and characterizing stratigraphic layers (Figure 2.2c). More information on the used devices is given in chapter 3.

## 2.4 Spectral imaging - fields of use

Spectral imaging and spectroscopy in general, nowadays, are used throughout many disciplines. The fields of use include, among others, agriculture, food processing, mineralogy, astronomy, chemical imaging, surveillance, and art conservation. Being one of the most comprehensive fields of use, remote sensing offers a variety of software solutions to deal with spectral data. Together with applications from the fields of signal processing and soil sciences, algorithms used for remote sensing offer a basis from which one can transfer concepts to the field of stratigraphic analysis or archaeological research.

**Remote Sensing** Spectral imaging techniques have a long tradition in remote sensing (Figure 2.3). The first airborne application dates back to the nineteenth century and in 1959 the first (spaceborne) satellite equipped with optical sensors was launched (Vanguard 2, United States Naval Research Laboratory). Since then, the analysis of optical data and spectral images of constantly improving quality is a fixed component of remote sensing. As spectral sensors become lighter and more affordable, the field of airborne remote sensing is constantly extending, culminating in the now widely used and highly effective unmanned drones which can be equipped with different measurement sensors. The fields of study involving optical remote sensing are manifold and range from crop management and vegetation monitoring to the analysis of urbanisation and the detection of mineral deposits.

The most common algorithms for the analysis of spectral image data in the field of remote sensing lead to a segmentation of the image data (i.e. image pixels) into different classes. Each pixel is assigned to a certain class, representing for example a specific land cover class (e.g. grassland, urban areas, or forests). Various studies, used this approach to detect changes in land cover over a defined timespan (e.g. Schubert et al., 2018) or to monitor and assess impacts of urbanisation (e.g.

Millward, 2011; Taubenböck et al., 2012). In remote sensing, optical data in many cases constitutes an additional layer of information and is therefore often combined with other data like digital elevation models (e.g. van Beijma et al., 2014) or radar data (e.g. Waske and Benediktsson, 2007; Joshi et al., 2016).

**Soil Sciences**    Spectral imaging is used in the soil sciences since at least the 1980s (e.g. Baumgardner et al., 1985). Alongside the improvement of spectral imaging techniques an increase in soil-related studies is visible (Figure 2.3). Numerous studies examined the spatial distribution of soils and soil characteristics as part of remote sensing approaches, mostly based on the integration of ground-proofed soil data (e.g. Hengl et al., 2007). Only in the last years an increase of studies that use spectral imaging for soil profile description, sediment analysis and the identification of diagnostic soil horizons can be observed (e.g. Udelhoven and Schütt, 2000; Steffens and Buddenbaum, 2013, 2012; Zhang and Hartemink, 2019c; Haburaj et al., 2019).
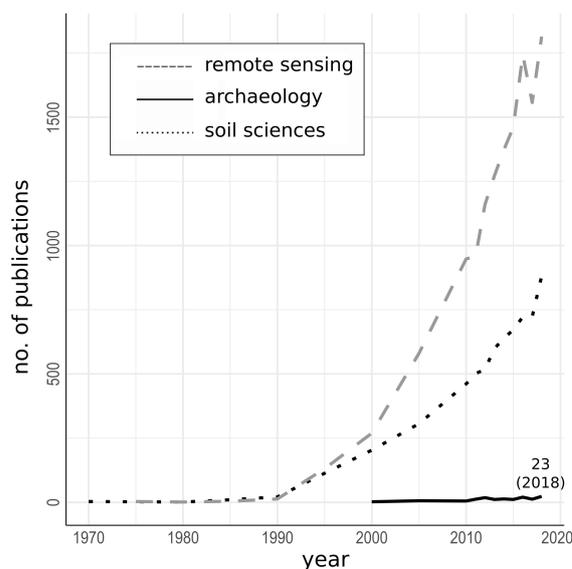


**Figure 2.3:** Number of articles covering a topic related to *spectral imaging* by year published and discipline, as registered at Web of Science (accessed August 2019, data range: 1970 to 2018).

Characteristic variances of the reflectance curves of soil and sediment layers allow for the quantitative analysis of various physical and chemical parameters. Most commonly, the examined parameters comprise organic matter and iron oxides (e.g. Stoner and Baumgardner, 1981; Hill and Schütt, 2000; Chen et al., 2020), as well as heavy metal concentrations (e.g. Kemper and Sommer, 2002; Vohland et al., 2009) and clay minerals (e.g. Altinbas et al., 2005). Recent studies utilised the relation of the spectral reflectance with these parameters to investigate, for example, the usage of RGB and hyperspectral imaging for the mapping of organic matter throughout soil profiles (e.g. Steffens and Buddenbaum, 2012; Hobley et al., 2018; Zhang and Hartemink, 2019a). Application of this methodology remains, however, mostly limited to laboratory studies. The works of Zhang and Hartemink (2019a), Haburaj et al. (2019) and Zhang and Hartemink (2019b) constitute experimental field studies: they successfully delineate stratigraphic layers of multiple soil profiles based on digital image data. Furthermore, the prediction of organic matter based on digital image data was successfully transferred to the microscale morphometric analysis of soils by Heil et al. (2020).

Quantitative colour measurements are sometimes used throughout the soil sciences to support stratigraphic documentation (e.g. Sprafke, 2016). A review of applications of spectroscopic measurements for this purpose is, for example, given by Debret et al. (2011), who conclude that spectroscopy allows for a "[...] fast, high resolution and non destructive analysis in the field [...]" (p.15). Fur-

thermore, Chen et al. (2019) used selective spectroscopic measurements and statistical modelling to determine soil classes based on VIS-NIR data.

**Archaeology**    Similar to remote sensing, the application of spectral imaging techniques in archaeology is strongly connected to air- or spaceborne archaeology. Up to now it is mainly used for the detection and mapping of archaeological sites. Several studies use hyperspectral or multispectral satellite images to improve the detection of archaeological crop marks (e.g. Lasaponara and Masini, 2007; Aqdus et al., 2012). Similar work was conducted by Winterbottom and Dawson (2005) who analyse multispectral airborne images to detect archaeological sites via soil marks, crop marks and small topographic variations. Comer and Harrower (2013) collected multiple case studies that also focus on the topic of site detection. All of the mentioned studies have in common that they act at a medium-sized or large spatial scale, often examining widespread geographical areas. Certainly, this is accounted for by the technical limitation of hyper- and multispectral imaging and the traditional focus on space- and airborne devices. In the recent past, the significant technological progress in the field of unmanned aircraft systems (UAS) led to an increased amount of archaeological excavations utilizing this new technology. In this context, the usage of optical sensors, which can be mounted to an UAS, becomes increasingly popular in archaeological projects. For example, a study conducted by Retzlaff (2015), which uses an multispectral camera system mounted to an UAS, also concentrates on crop and soil mark analysis for archaeological research purposes.

In the context of archaeological projects, quantitative colour measurements, based on spectral recordings or systematic RGB data, are mostly concerned with the documentation of artifacts (e.g. Bratisi et al., 2018), while on-site documentation of sediment colour still heavily relies on logging standards like the Munsell soil colour chart.

## 2.5    Image analysis algorithms

There are multiple statistical algorithms used to analyse image data throughout this thesis. In the following, a short summary of relevant terminologies and methods is given. A summary of the algorithms used in this thesis is given in chapter 3.

**Supervised versus unsupervised learning**    The two main types of machine learning algorithms which are used for the statistical analysis of image data are unsupervised learning algorithms (e.g. cluster analysis) and supervised learning algorithms (e.g. regression analysis, decision trees, support vector machines). The main difference between these two approaches is that supervised learning requires the user to define classes by which the data should be classified, while unsupervised learning defines these classes by its own. The latter process leads to statistically coherent classes, which, however, require the user to interpret them after the classification was performed. Supervised classification, on the other hand, is highly dependent of the quality of the initial input. A third type of algorithms called reinforcement learning analyses the data sequentially by given positive and negative rules.

Regarding *supervised learning* techniques, there is a large number of algorithms which were presented and applied in the last decades. Recent studies in the field of remote sensing show a high potential of the *random forest* classifier developed by Breiman (2001) (e.g. Stefanski et al., 2013; Hengl et al., 2015). Other studies report good performance of, for example, support vector machines (e.g. Waske and Benediktsson, 2007) or penalised linear discriminant analysis (e.g. Goetz et al., 2015). In their review of various algorithms, Goetz et al. (2015) argue that, while random forest and bundling classification (in their case study) slightly outperform other techniques, the differences between all the examined modern classifiers are non-significant. The quality of the produced classifications, however, may vary: for example, according to Steger et al. (2015) some algorithms produce more homogeneous classes than others. When applying these supervised learning algorithms to spatial data, researchers should always be aware of the problems of spatial auto-correlation and thereby overfitting; one way to prevent this is to use cross-validation techniques and bootstrap strategies (Brenning, 2012).

In contrast to the above described methods, *unsupervised learning* techniques do not require the definition of classes or training data and thus assumptions prior to the performed classification. One simple yet often applied approach is the k-means clustering algorithm, which divides the image data into $k$ different clusters (or classes) based on either the complete data or a subset of it. Instead of a strict assignment to a class, also a membership to a class can be calculated; this so-called *fuzzy* approach thereby allows the allocation of image data (or pixels) to multiple classes. Cai et al. (2007) introduced more advanced forms of such a fuzzy algorithm: the so called fast generalised fuzzy c-means algorithms. By the inclusion of spatial constraints they were able to produce results that are evidently less susceptible to image noise than the traditional fuzzy c-means. Clustering algorithms such as k-means or c-means were successfully used for the analysis of soil data by, for example, Steffens et al. (2014) or more recently Zhang and Hartemink (2019a).

**Pixel-based versus object-based algorithms**  Hyperspectral sensors capture between 20 an 2500 different channels for each scene. In combination with a high spatial resolution or extent, this can lead to a very high amount of data. Many algorithms used for image analysis are based on single pixels. These *pixel-based* approaches treat every image pixel as an individual measurement and therefore ignore values of adjacent pixels. The advantage of such an approach is that all available data can be included in the performed analysis. However, this high amount of data and the included time-consuming and resource-heavy processing often leads to tradeoffs in the form of (i) an active decreasing of the spatial resolution, (ii) a dimension-reduction of the data by calculating derivatives like principal components (PCA, see Hotelling, 1933), or (iii) the limitation of the analysed data to a subsample with later extrapolation of the results. All image analyses included in this thesis utilise pixel-based approaches (see chapter 3).

A different approach called *object-based* uses localised groups of pixels instead of single pixels. Such groups are often called superpixels. This method is often used to segment image data into classes (e.g. Blaschke, 2010). Various algorithms can be used for this: Stefanski et al. (2013) propose the usage of contour-relaxed superpixels (CRS, see Conrad et al., 2013) in combination with a random forest algorithm to be highly accurate when mapping land cover. Other techniques include

for example the watershed algorithm (Meyer, 1992, e.g.) or simple linear iterative clustering (SLIC, see Achanta et al., 2012). A comprehensive comparison of superpixel algorithms was carried out by Stutz et al. (2018) via a benchmark which allowed the evaluation and ranking of numerous algorithms. They attribute superior performance to ETPS (Yao et al., 2015), SEEDS (Bergh et al., 2015), ERS (Liu et al., 2011), CRS (Conrad et al., 2013), ERGC (Buyssens et al., 2014) and SLIC (Achanta et al., 2012) algorithms.

## Material and methods

## 3.1 Profile locations

Throughout the presented studies archaeological and geoscientific cross-sections from three different locations are examined (Figure 3.1). Two soil profiles, located in the area of a historical vineyard near Chorin (Brandenburg, Germany), were documented and analysed in chapter 4. The archaeological profile examined in chapter 5 is part of an archaeological trench, situated at the flank of a Bronze Age burial mound near Seddin (Brandenburg, Germany). Additionally, two archaeological profiles from a trench located in the vicinity of the Great Temple of Yeha (Tigray, Ethiopia) were examined in chapter 6.

To ensure reliable stratigraphic interpretations of the respective site, experts, who are familiar with the local site formation processes and the stratigraphic record, were involved in the presented studies. Cross-sections with an existing and profound stratigraphic documentation produced by these experts were chosen, allowing for the evaluation of the performance of the statistical analyses carried out in the single studies against expert knowledge.

## 3.2 Spectroscopic devices used

A basic form of spectral imaging can be found in every modern digital camera. Given access to the raw image files produced by common RGB camera systems, one has access to three distinct image bands, covering specific regions of the blue, green and red part of the visible part of the electromagnetic spectrum. While the term 'spectral imaging' is usually related to more advanced spectral sensors, modern RGB cameras must therefore also be regarded as multispectral devices.

Several studies found promising results when analysing soil and sediment data on the basis of the limited spectral capabilities of regular RGB cameras (e.g. Zhang and Hartemink, 2019a; Haburaj
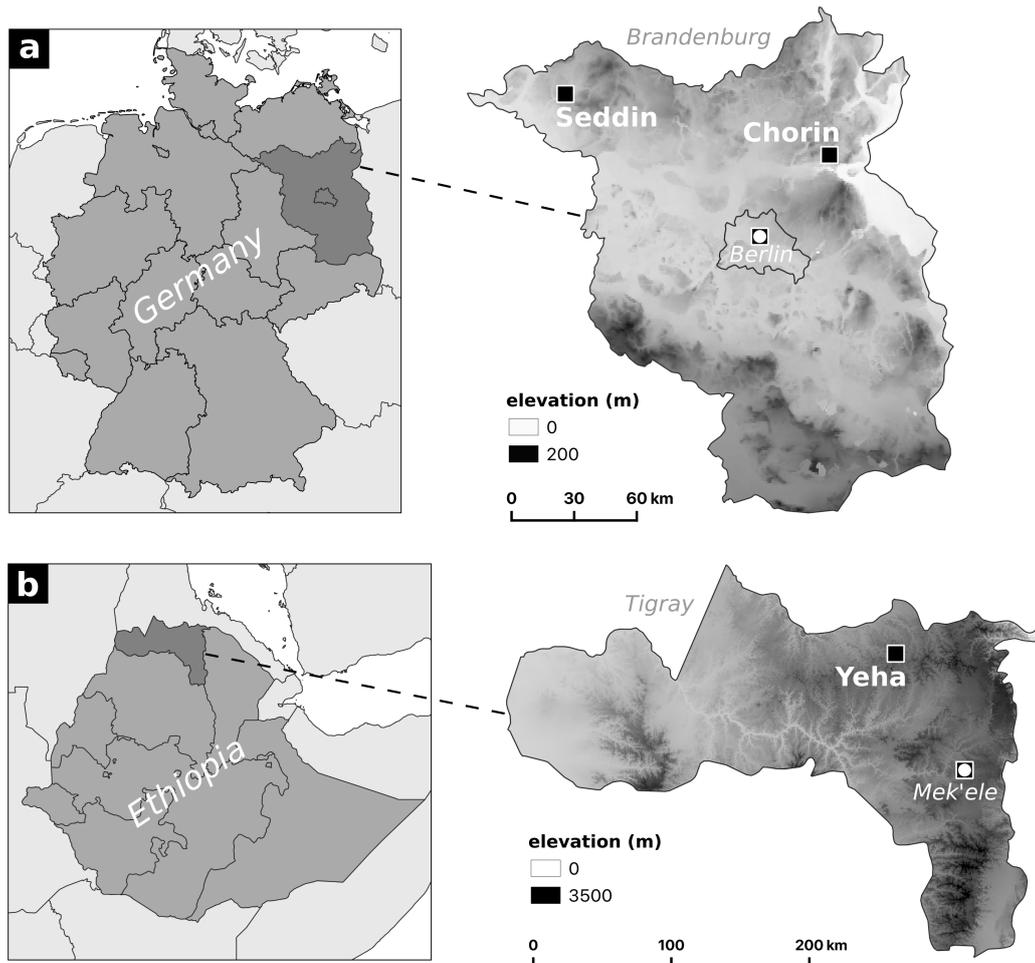
**Figure 3.1:** Locations of the profiles examined. (**a**) The two soil profiles examined in chapter 4 are located near Chorin; one archaeological profile near Seddin (chapter 5). (**b**) Two archaeological profiles belong to an archaeological trench in Yeha (chapter 6). Elevation data: GTOPO30 data (U.S. Geological Survey).

et al., 2019; Heil et al., 2020; Haburaj et al., 2020a). Since these devices are far more common in fieldwork equipment than specific multi- or hyperspectral sensors and therefore hold much potential, all studies at hand include data that was acquired with a traditional RGB camera. To ensure comparability of the image data used throughout the presented studies, the spectral properties of this RGB camera (Sony ILCE-6000) were recorded by applying a systematic measurement of monochromatic light, emitted from a Newport MS257 monochromator using a halogen light source. From the resulting data we were able to derive the central wavelengths of the red, green, and blue channels as well as the respective FWHM (full width at half maximum) values and thus the approximate bandwidth (Figures 3.2, 3.3). More details regarding the spectral properties of the RGB camera used are given in chapter 4.

More advanced multispectral sensors are able to capture very specific wavelengths of the electromagnetic spectrum, often resulting in more image bands than the three bands captured by RGB cameras. Beyond, hyperspectral sensors offer the most detailed capturing method of spectral information by capturing a very high number of equidistant image bands and most commonly allowing one to analyse continuous spectra in a specific spectral range.
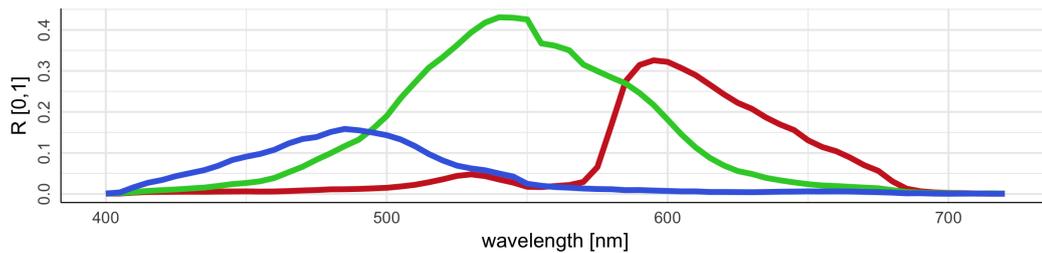
**Figure 3.2:** Spectral sensitivity of the RGB sensor used (Sony ILCE-6000), measured using monochromatic light from a monochromator. Spectral reflectance (R) normalised to [0,1].

There is a variety of techniques available when collecting spectral information with spectral imaging devices. The most commonly used capturing techniques are (i) point scanning or whiskbroom mode, (ii) line scanning or pushbroom mode, (iii) plane or area scanning, or (iv) single shot or snapshot mode (Wu and Sun, 2013; Adão et al., 2017). Most large-scale air- or spaceborne spectral imaging devices use either optoelectronical pushbroom scanning (e.g. SPOT, MODIS, LANDSAT 8, see Shaw and Burke, 2003) or optomechanical whiskbroom scanning (e.g. LANDSAT 7) to capture spectral bands which hold information which is related to various parameters of land and water surfaces.

Technological advancements in the last decade led to smaller and lighter spectral imaging devices which become increasingly popular when unmanned aerial vehicles (UAVs) are used for remote sensing (Robbins, 2018). A review by Adão et al. (2017) shows that pushbroom scanners are dominating the market of spectral sensors which can be attached to UAVs. The comparatively faster and often lighter snapshot systems are a more recent development and only available from a limited number of manufacturers. Generally speaking, a higher spatial and spectral resolution of the sensor is often related to a much higher weight and cost of the device (Adão et al., 2017).

Snapshot-style hyperspectral cameras gained popularity in the last years due to their low weight, fast and easy image acquisition and often low power consumption. These characteristics render them very suitable for remote sensing, often conducted with small UAVs. The same characteristics are advantageous for many terrestrial applications of spectral imaging, including plant monitoring (Behmann et al., 2018), fruit quality assessment (Rungpichayapichet et al., 2017) or proximal soil sensing (Jung et al., 2015).

Working in the context of terrestrial archaeological and geoscientific projects, we chose spectral devices that include specific characteristics regarding their weight, size and power consumption. The examined study areas can be described as remote places with limited power supply. Additionally, the on-site situation during archaeological or geoscientific excavations often includes very limited space for setting up measurement devices and measurements often have to be conducted fast since fieldwork schedules often are tight. Therefore, two snapshot-style spectral cameras with a comparably high spatial resolution are used throughout the presented studies (Table 3.1). A multispectral snapshot camera with six spectral bands and a hyperspectral snapshot camera with 40 spectral bands were chosen. This setup allows for the capturing of soil and sediment profiles in a high spatial resolution and likewise helps to assess the importance of the factors spectral range and spectral resolution for the topics examined.
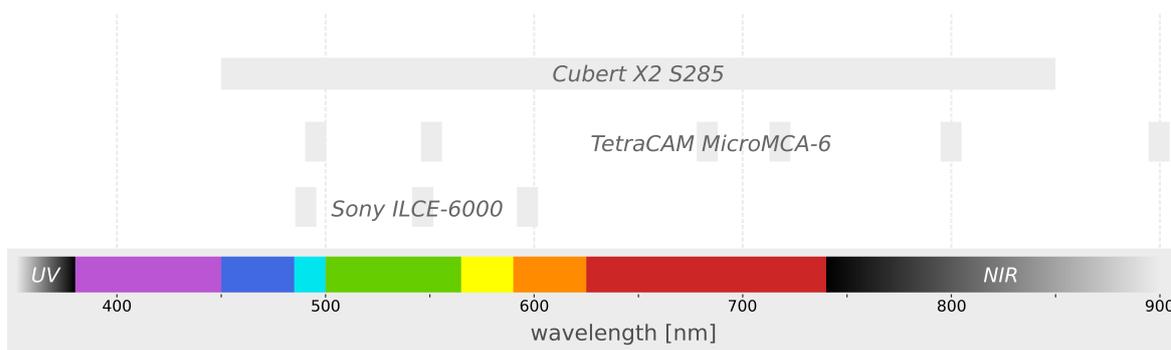
**Figure 3.3:** Spectral range and resolution of the three camera systems used throughout the presented studies. See Table 3.1 for more details.

**Table 3.1:** Technical specifications of the camera systems used. Central wavelengths of the spectral channels blue, green, red, red edge, NIR1, and NIR2 of the used Sony ILCE-6000 and TetraCAM MicroMCA-6. The Cubert X2 S258 captures the range from 450 to 850 nm in 10 nm steps.

|  | **Sony ILCE-6000** | **TetraCAM MicroMCA-6** | **Cubert X2 S258** |
|---|---|---|---|
| B | 485 nm | 490 nm | 450 nm |
| G | 540 nm | 550 nm | … |
| R | 595 nm | 680 nm | … |
| red edge | - | 720 nm | … |
| NIR1 | - | 800 nm | … |
| NIR2 | - | 900 nm | 850 nm |
| spatial resolution | 5696 × 4272 px | 1280 × 1024 px | 512 × 272 px |
| colour depth | 8 bit | 10 bit | 16 bit |

The two snapshot-style spectral imaging sensors, used in the presented studies, cover the VIS and NIR part of the electromagnetic spectrum (Figure 3.2). The multispectral device TetraCAM MicroMCA-6 captures image bands at 490, 550, 680, 720, 800, and 900 nm. The hyperspectral sensors of the Cubert X2 S258 cover the range from 450 to 850 nm in 10 nm steps, totalling in 40 equidistant image bands.

Two of the conducted studies (chapters 5 and 6) additionally feature an ASD FieldSpec II spectroradiometer, covering the range from 350 to 2500 nm in 1 nm resolution. This device was used to more accurately measure the spectral reflectance of the sediment samples examined in these studies. Contrary to the spectral camera systems described above, the used spectroradiometer is not able to capture spectral images but only measures the spectral reflectance of a selected spot of an object. Readings of the spectroradiometer were used for validation of the data acquired with the spectral cameras (chapter 6) and to determine the potential of a higher spectral range and resolution when compared to the multi- and hyperspectral camera sensors used throughout the studies conducted (chapter 5).
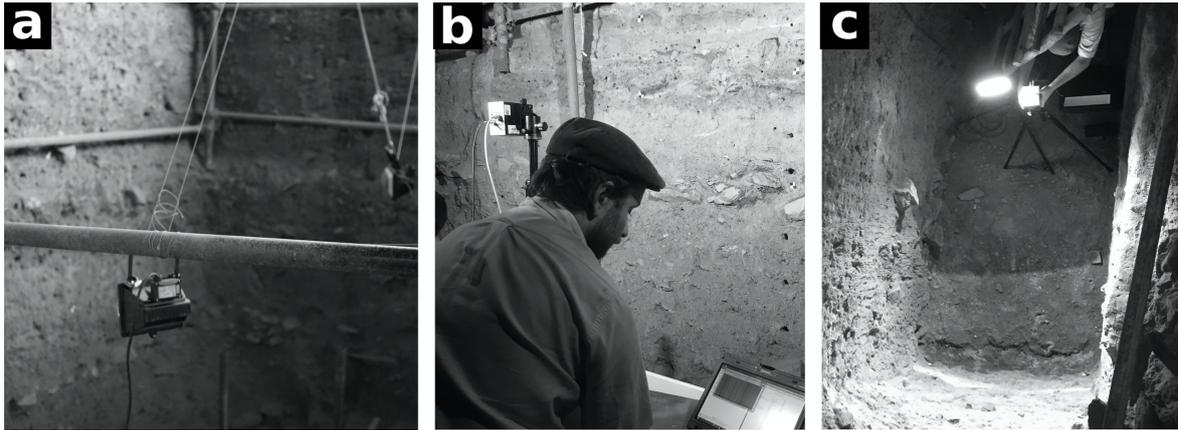
**Figure 3.4:** Imaging setup during fieldwork. Hyperspectral images, used throughout chapters 4 and 6, were captured using artificial light, which was positioned to create uniform lighting (a). Image acquisition required the hyperspectral camera (Cubert X2 S258) to have a constant wired connection to the imaging software (b). In tight spaces, the usage of a single artificial light source can be necessary (c).

## 3.3   Image acquisition

The five profiles examined throughout the presented studies were recorded with different combinations of the above mentioned spectral cameras. The camera selection and image acquisition setup of each study were adapted to the respective scope of the conducted research.

To evaluate the potential of spectral imaging for the delineation of stratigraphic layers, the two soil profiles included in chapter 4 (Haburaj et al., 2019) were captured using a RGB camera (Sony ILCE-6000) and a hyperspectral snapshot camera (Cubert X2 S258). Each profile was captured by a single recording of each camera and overlying image pairs for each profile were produced for further processing. By including the significant differences of the two camera systems used (Table 3.1), we were able to assess the impact of spatial resolution, spectral resolution and spectral range on the conducted image classifications.

Taking into account the good performance of the RGB camera used in chapter 4 (Haburaj et al., 2019), we compared image classification results produced from data recorded with the RGB camera Sony ILCE-6000 and a multispectral camera (TetraCAM MicroMCA-6) in chapter 5 (Haburaj et al., 2020a). The archaeological profile examined was recorded in a similar manner as the previously documented soil profiles: single recordings of the profile were captured using two different camera systems, producing overlying image pairs of the two datasets.

The results presented in chapter 4 (Haburaj et al., 2019) and chapter 5 (Haburaj et al., 2020a), indicate that spectral images, which would include spectral data of similar quality as high-resolution (selective) spectral recordings produced by an ASD FieldSpec II spectroradiometer, could prove highly effective for delineating stratigraphic layers. Additionally, findings of, for example, Zhang and Hartemink (2019a) showed promising results for mapping sediment properties spatially based on image data. For these reasons, we examined the performance of the hyperspectral camera Cubert X2 S258 in more detail in chapter 6 (Haburaj et al., 2020b). The image aquisition of the two examined archaeological profiles significantly differs from the two previous studies: while a single RGB photo of each profile was captured using the RGB camera Sony ILCE-6000, multiple images

were captured using the hyperspectral camera Cubert X2 S258 and then manually combined to a hyperspectral orthophoto, allowing us to increase the spatial resolution of the hyperspectral image data. On-site quality control of the hyperspectral recordings was performed using the software UHD examineR by Haburaj (2019). The produced hyperspectral orthophotos and the RGB images were, again, used to create overlapping image pairs for further processing.

### 3.3.1   Lighting and white references

One major drawback of optical data is that it is highly dependent on lighting conditions. This topic was therefore handled with care throughout the presented studies. Contrary to most studies in the past, the presented work was not carried out in the laboratory but during fieldwork. Lighting conditions therefore had to be adjusted according to the respective on-site situation. To counteract over-illumination or the casting of shadows, halogen light sources were used in chapters 4 and 6. The profiles examined in chapter 4 were captured illuminated by a mixture of natural daylight and artificial light. Recordings for chapter 6 were captured after nightfall and thus illuminated by halogen light only (Figure 3.4). The latter procedure eliminates the possible influence of slightly changing lighting conditions between single measurements due to small clouds and scattered light of adjacent objects. The profile examined in chapter 5 was illuminated nearly uniformly by natural daylight. During processing of the data, minor differences in image brightness could successfully be compensated for by combining spectral slope rasters with image data that was transformed to the CIE-LAB colour space (see chapter 5).
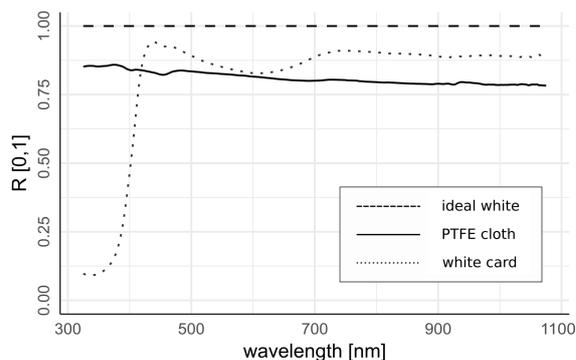


**Figure 3.5:** Comparison of white reference materials used throughout the presented studies. An uncalibrated SRM-990 Spectralon white reference was used as ideal white. Spectral reflectance (R) normalised to [0,1].

The internal setup of the Cubert X2 S258 requires the definition of a white reference prior to recording hyperspectral images. The spectra of the different white references used in the presented studies are depicted in Figure 3.5. The shown spectra were recorded using an ASD FieldSpec Handheld 2 Spectroradiometer (325 - 1075 nm) and processed using a Savitzky-Golay filter (second order polynomial, 21 bands) from the *signal* package (signal developers, 2014). In chapter 4 a standard white card was used as it is used by professional photographers and shipped with the Cubert X2 S258. To increase the quality of the spectral recordings, later measurements (chapter 6) were captured using a polytetrafluoroethylene (PTFE) coated cloth as white reference. PTFE is also used by leading manufacturers in the field of optical sensor systems for producing highly accurate white references with a nearly ideal lambertian, diffuse reflectance throughout the VIS-NIR. When compared to the ideal white of a SRM-990 Spectralon uncalibrated white reflectance standard (approx. 99 % diffuse reflectance),

the PTFE coated cloth shows only minor drawbacks: it introduces some noise but creates a nearly constant reflection up to 1075 nm (Figure 3.5). The signal of the PTFE coated cloth, however, is much more stable than the one of the white card.

## 3.4  Laboratory sediment analyses

The image classifications performed in chapter 4 were carried out to assess the value of semi-automatic image processing for the delineation of stratigraphic layers. Quality of the results was evaluated by a comparison of the image-based stratigraphic delineation to the delineation based on expert knowledge and laboratory analyses carried out by these experts. Those laboratory analyses, carried out by Böttcher (2002), include, among others, the determination of grain size composition, soil organic carbon, pH value and the measurement of element concentrations. No further sediment analyses were conducted as part of the study presented in chapter 4. Rather, the focus of Haburaj et al. (2019) lies on image acquisistion and processing.

Including a direct comparison of clustering results of spectral data on the one hand and more traditional laboratory sediment analyses on the other hand, chapter 5 (Haburaj et al., 2020a) involves the results of multiple sediment analyses carried out in the laboratory. The captured multispectral images were complemented by spectral measurements which were carried out in the laboratory. An ASD FieldSpec II (350 to 2500 nm in 1 nm steps) was used to capture the spectral reflectance of the extracted sediment samples (Figure 3.6). A TruSpec CHN analyser by Leco was used to determine the total carbon content of the samples. As none of the samples showed signs of carbonate when treated with 9.9% HCl acid, the measured total carbon content is interpreted as soil organic carbon (SOC). The pH values of the samples were measured in a 1:2.5 solution of 10 g of dried sediment and 25 mL of 0.01 M KCl using a handheld pH meter (Hanna Instruments).

Some sediment parameters were determined by using the same methods in chapters 5 and 6. Grain size distribution in both cases was determined with a laser diffraction particle size analyser LS13 320 by Beckman-Coulter. Classification of the data was carried out following Ad-hoc Arbeitsgruppe Boden (2005). Water content of the extracted samples was determined gravimetrically. Samples were weighed before and after drying at 105 °C. The water content was calculated according to Blume et al. (2011) and reported in mass%. Additionally, element contents of the samples included in chapters 5 and 6 were measured using a Thermo Scientific Niton XL3t portable energy-dispersive X-ray fluorescence spectrometer (p-ED-XRF). Internal pa-
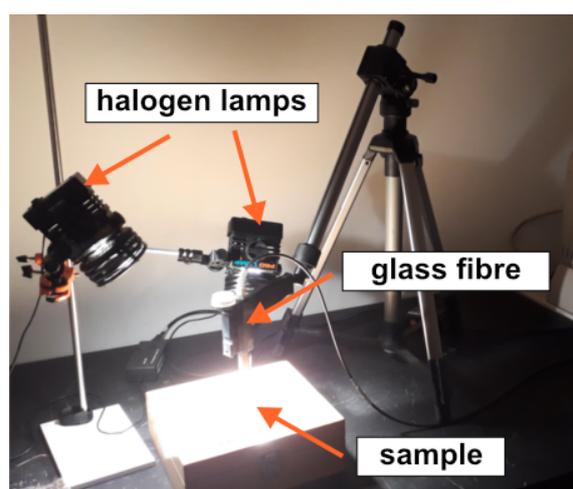


**Figure 3.6:** Laboratory setup used to capture the spectral reflectance of sediment samples with an ASD FieldSpec II (350 to 2500 nm).

rameter calibration was set to 'mining mode' and all measurements were recorded in %. Quality control was conducted using certified reference materials (CRMs). Measurements with a value smaller than or equal to three times the $1 - \sigma$ error were discarded. Elements were only processed if the number of values below this limit of detection did not exceed 10% of the respective measurements.

Since the analyses described in chapter 6 do heavily rely on the sedimentological record of the two archaeological profiles examined, additional laboratory analyses were carried out to offer a more complete record. Complementing the p-ED-XRF measurements, we additionally included the measurement of element contents using a PerkinElmer Optima 2100 DV inductively coupled plasma optical emission spectrometer (ICP-OES) after aqua regia digestion. This allowed us to reliably use element concentrations of Si, K, Mg and Ca measured by p-ED-XRF, and Na, Fe, Al, Ti and $PO_4$ concentrations as provided by ICP-OES data for further analyses. Carbon measurements slightly differ between chapters 5 and 6: the samples analysed in chapter 6 do include carbonates, which is why total inorganic carbon (TIC) was additionally measured using a Woesthoff Carmhograph C-16. SOC was then calculated by subtracting TIC from the total carbon contents measured with a TruSpec CHN analyser by Leco. In chapter 5, SOC was regarded equal to total carbon contents. SOC values used in chapter 6 were additionally transformed to soil organic matter (SOM), using a factor of 2, according to Pribyl (2010). Involving the analysis of iron oxides, the laboratory analyses performed in chapter 6 do also include the measurement of the mineral conposition of the samples. This was conducted using a Rigaku MiniFlex 600 X-ray powder diffractometer (XRD) with a copper $k\alpha$ tube. The element concentrations measured by p-ED-XRF and ICP-OES and discussed in chapter 6 were used to calculate the following weathering indices: Ca-Ti-ratio (CTR; Bétard, 2012), the Ruxton-ratio (RX; Ruxton, 1968), the Sesquioxide-ratio (SQ; Ruxton, 1968), the K-Ti-ratio (KT; e.g. Davies et al., 2015) and Parker's index (PI; Parker, 1970).

## 3.5 Data processing

### 3.5.1 Image alignment

Comparing statistical results based on the recordings of different imaging devices requires the correct alignment of single images acquired during fieldwork. The hyperspectral image data recorded with the Cubert X2 S258 and analysed in chapter 6 are a composite of multiple hyperspectral images: we created hyperspectral orthophotos by manually combining multiple images. We did so by identifying overlapping geometries in the images and using black-and-white reference targets. This step was carrid out in the software QGIS (v3.4) using thin plate spline (TPS) transformation and nearest neighbour (NN) resampling and including the rectified RGB images as spatial reference. To eliminate minor differences between single images histogram matching was used (R package RStoolbox, Leutner and Horning, 2017). The processed images were merged using GDAL (v2.4.1).

A similar process was included in the workflows presented in chapters 4, 5 and 6 to generate overlapping image datasets to allow for the comparison of the results produced by the different cam-

**Figure 3.7:** Spatial filters applied to a exemplary pixel image (left). The median filter used throughout chapters 4, 5 and 6 eliminates smaller image noise while preserving main features.

era systems: the multispectral images (chapter 5) and the hyperspectral images (chapters 4 and 6) were spatially adjusted in QGIS (v3.4) using TPS transformation and NN resampling to match the respective rectified RGB images.

## 3.5.2  Image pre-processing

Image data acquired during fieldwork often includes the irregular casting of small shadows due to sediment texture. Additionally, most images acquired with multi- and hyperspectral cameras contain low to high amounts of image noise. These interfering factors can heavily influence the results of statistical image analyses in general. A common practice to counteract these disturbances and apply noise reduction is the application of spatial filters. Today, numerous filter algorithms are used successfully in various fields of research (e.g. Dravida et al., 1984; Hoshyar et al., 2014). Figure 3.7 shows two of the most commonly used algorithms: the moving mean and moving median filter. Both filters can be applied to image data by including a moving window of a defined size (x,y). As visible in Figure 3.7, the mean filter blurs the original image, while the median filter removes image noise and preserves edges. The studies presented in chapters 4, 5 and 6 do include the usage of median filters as the intended effect of the applied image filters is to remove noise while preserving the measured spectral reflectance of the soils and sediments as accurately as possible. Filtering was carried out using the raster package for R (Hijmans, 2019).

The hyperspectral camera used in chapters 4 and 6 (Cubert X2 S258) produces continuous spectra for each image pixel by offering an equidistant spectral imaging setup. As discussed in those chapters, the spectral information acquired with this camera is prone to noise in the spatial domain as well as the spectral domain. The recorded hyperspectral image data was therefore additionally filtered using signal processing algorithms in the spectral domain. Each spectrum (i.e. each pixel) was filtered using a Savitzky-Golay filter of second polynomial order over a defined number of wavelengths (Savitzky and Golay, 1964; Lehnert et al., 2017). In chapter 6 the spectral data between 560 and 670 nm additionally was interpolated using piecewise cubic hermite interpolating polynomial interpolation (R package signal, signal developers, 2014). This interpolation was necessary to mask the high noise produced by the automated adjustment between the two overlapping image sensors implemented in the Cubert X2 S258.

Throughout the presented studies we created multiple derivatives of the image datasets to assess their influence on the applied analyses. Multiple authors suggest the CIELAB colour space (Commission Internationale de l'Eclairage) to be highly suited for quantitative analyses (e.g. Melville and Atkinson, 1985; Viscarra Rossel et al., 2006; Jarmer et al., 2009; Hill et al., 2010; Sprafke, 2016;

Vodyanitskii and Kirillova, 2016; Zhang and Hartemink, 2019a). The CIELAB colour space consists of three image bands: the band L* holds information on the lightness of the image, while the two bands a* and b* capture chromaticity (red/green, yellow/blue). We therefore included colour transformation to the CIELAB colour space in chapters 5 and 6. The used RGB data and selected bands of the multispectral data were converted to CIELAB data using the convertColor() function for R. As several sediment properties influence spectral reflectance over a specific spectral range (e.g. Baumgardner et al., 1985), we additionally calculated the slope between the single image bands of the RGB and multispectral data used in chapters 4 and 5. Furthermore, to assess the influence of reducing and expanding the data dimensionality on the clustering results, principal component analysis (PCA, calculated after Leutner and Horning, 2017) was performed based on the RGB and multi- and hyperspectral images (chapters 4 and 5) and minimum noise fraction (MNF, calculated after Pebesma, 2012) was calculated for the hyperspectral images in chapter 4.

### 3.5.3   Image classification

Unsupervised image classification was used throughout chapters 4 and 5 to semi-automatically delineate stratigraphic layers based on the acquired image data. The k-means clustering algorithm after Hartigan and Wong (1979) was used due to its simplicity and low computational effort. The clustering was carried out on a subset of each dataset and afterwards used for prediction via the clue package for R (Hornik, 2005). In both chapters, different datasets were created and used as input for the performed classification. This approach allowed us to assess the positive and negative influence of specific pre-processing steps on the overall quality of the classification results.

### 3.5.4   Regression analysis based on image data

Spectral reflectance of an object is highly dependent of its chemical and physical composition (e.g. Baumgardner et al., 1985). Due to this relation, spectral recordings are often correlated with soil and sediment properties using statistical models, resulting in maps of the examined properties (e.g. Hill et al., 2010; Steffens and Buddenbaum, 2012; Haijun et al., 2017; Hobley et al., 2018). Soil organic matter (SOM, also SOC [soil organic carbon]) shows a strong connection with the visible and near-infrared spectral reflectance and is therefore frequently investigated (e.g. Baumgardner et al., 1985; Hill and Schütt, 2000; Mutuo et al., 2006; Brown et al., 2006; Sankey et al., 2008; Volkan Bilgili et al., 2010; Steffens et al., 2014; Hobley et al., 2018). In chapter 5 we created a map of SOM throughout the examined profile by using ordinary least squares regression (OLS) on the multispectral image data and the SOM values obtiled from sediment samples. As such parameter maps potentially contribute to the general understanding of stratigraphies and likewise render layer delineation more traceable, this topic was examined further in chapter 6. Here, we used random forest regression (Breiman, 2001; Cutler et al., 2018) to create parameter maps of SOM, hematite, calcite, multiple weathering indices and grain size composition based on RGB, RGB derived CILAB and hyperspectral image data. Modelling of the regressions and predictions based on the image data were carried out using the R programming language.

### 3.5.5   Processing of sedimentological data

Element concentrations of the sediment samples obtained from p-ED-XRF and ICP-OES and examined in chapters 5 and 6 were controlled for their quality using certified reference materials (CRMs). The CRMS in chapter 5 were only used to validate the measurements. In chapter 6 we used linear models to correct the measurements of the sediment samples based on the measured and actual values available for the measured CRMs. This was carried out using the lm() function for the R programming language.

In chapter 5 we compare clustering results of different selectively sampled data: p-ED-XRF measurements, ASD FieldSpec II measurements and selected pixel values of image data. Prior to clustering, the p-ED-XRF data was processed using a compositional data approach to manage the consistent sum-constrained model (Aitchison, 1986). The data was transformed using centered log-ratio (clr; Aitchison, 1986) and standardised using a robust z-transformation. The processed p-ED-XRF data as well as the other datasets were afterwards clustered using hierarchical clustering. Additionally, the usage of depth-constrained hierarchical clustering (Gill et al., 1993) using CONISS (Grimm, 1987) was examined.

CHAPTER 4

# Delineating stratigraphic layers based on image data

## 4.1 Introduction

Methods for documenting archaeological remains are highly linked to technological progress (see chapter 2). In the recent past, computer aided documentation (e.g. CAD, GIS, digital photography) became widely accepted as a standard procedure for excavations. Carried out correctly, digital photography is more than a form of visual documentation: it is a physical measurement of spectral reflectance. The continual technological progress of digital photography in general (i.e. higher spatial resolution, better low-light performance, higher portability, better sharpness) is recently accompanied by advances in the field of hyperspectral imaging. This development led to the production of small, and inexpensive sensors, which can easily be used during fieldwork and capture the visible and near infrared part of the electromagnetic spectrum and therefore capture more information than common RGB photographs.

In the following study, the potential of high resolution RGB imaging and portable hyperspectral imaging for the documentation of archaeological profiles is assessed. It is examined, (i) if diagnostic soil and sediment layers can be delineated semi-automatically by performing a statistical classification of the image data, (ii) if such image data can be reliably acquired under fieldwork conditions, and (iii) what influence the different technological specifications of the used cameras (spectral resolution and range as well as spatial resolution) have on the classification results. The results of this study help to assess the potential of digital imaging for stratigraphic documentation and interpretation.

## 4.2 Paper 1: Evaluating the potential of semi-automated image analysis for delimiting soil and sediment layers

### Abstract

Established methods for delineating anthropogenic and natural strata during fieldwork are based on the visual and tactile perception of excavators. Modern image analysis techniques can help to ensure objectivity and reproducibility when documenting sections and plana. Within this study we examine the unsupervised classification of digital images as a technique for delimiting layers and identifying stratigraphic features. Assessing the potential of this approach, we exemplarily captured soil profiles with high-contrast stratigraphy, located in the area of a historical vineyard (Brandenburg, Germany). Reproducible analyses were carried out using open-source software, allowing for the future advancement of the methodology utilized and providing a basis for the analysis of more complex stratigraphic sequences. We compare clustering results of high-resolution RGB and hyperspectral images (470-830 nm, 37 bands). Multiple pre-processing and processing steps are carried out to evaluate their influence. Our results render the semi-automatic analysis of RGB images helpful for stratigraphic interpretation.

**Keywords:** fieldwork; digital archaeology; unsupervised classification; stratigraphy; RGB imaging; spectral imaging; landscape archaeology

### 4.2.1 Introduction

Stratigraphic analysis, with its origin in geological sciences, is also a common part of geographical and archaeological fieldwork and mode of operation (e.g. Daniel, 1976). The human *reading* of geoscientific and archaeological sediment layers and sections under fieldwork conditions nevertheless relies heavily on professional expertise, including both technical and theoretical knowledge. State of the art documentation techniques are still based on what, for example, Wheeler (1960) suggested including in situ characterization of the soil texture (nowadays loosely based on the USDA soil texture triangle), technical drawings of sections and plana, and the numbering of layers and interfaces. In the recent past these procedures have been complemented by digital tools like digital photography and spectral color analysis. Although these techniques are widely accepted, assistive quantitative methods have emerged over the past decades, supplementing fieldwork results and thereby leading to the better and more transparent documentation of stratigraphic units (e.g. Bebermeier et al., 2018; Inglis et al., 2018). In most cases, these methods, which often include extensive work in the laboratory, are only applicable after in situ fieldwork is complete. As a consequence, the initial identification of stratigraphic units within archaeological or geoscientific sections during fieldwork remains a process that is of central importance, and at the same time relies on the experience and expertise of the recording scientist or technician. Logging standards such as compiled in the Bodenkundliche Kartieranleitung in Germany (Ad-hoc Arbeitsgruppe, 2005) provide support while recording strata in outcrops, but since identification depends on the visual and tactile perception of the respective excavator, it lacks reproducibility to a certain extent-and remains highly time consuming.

Here, we want to lay a basis for further development of an automated delineation of stratigraphic layers in geomorphological and archaeological outcrops using spectral reflectance. We are ultimately aiming for a fast and applicable tool that supports the recording of the stratigraphy of a site, in order to provide reproducible and reliable stratigraphic documentation that can be compiled instantly. Such a tool would be of high value especially under conditions with a high work load, such as rescue excavations. In the first step outlined in this paper we prove the general suitability of spectral reflectance for the delineation of a reconstructable and reliable stratigraphy. For this purpose, we selected some well-documented soil profiles with distinct layer boundaries from the Educational Soil Trail Chorin-Weinberg (Umweltbundesamt, 2010). Based on the outcomes discussed in this paper, we plan advanced experiments that will include outcrops and profiles with less apparent stratigraphic layers than those presented in this paper.

We want to examine whether semi-automated image analysis as applied by default to remote-sensing data, as well as data from laboratory spectral imaging, can be used to delimit layers from field outcrops, and thus whether it can support layer delineation of cross-sections under field conditions. We compare the performance of a digital RGB camera and a hyperspectral snapshot camera; we will demonstrate how semi-automated image analysis can be utilized to delimit stratigraphic layers of high-contrast archaeological and geographical profiles via unsupervised classification. We evaluate the potential of this approach for fieldwork purposes. We do not suggest that these analyses should replace traditional methods of stratigraphic determination; rather, the results should be understood as an additional layer of documentation that supports the potentially subjective perception of the expert at work with a physical measurement of spectral reflectance. While we solely focus on the act of layer identification, we likewise create a basis for further research regarding more complex stratigraphy and the characterization of the respective layers via the spectral signal. Contrary to Steffens and Buddenbaum (2013), for example, our work is carried out in situ, and is therefore the first study to apply hyperspectral snapshot imaging for archaeological and geoscientific fieldwork purposes. Our experimental design is based on the soil profiles of the *Bodenlehrpfad Chorin Weinberg* (or, the "Educational Soil Trail Chorin-Weinberg") (Umweltbundesamt, 2010), whose high-contrast stratigraphy offers a suitable basis for the evaluation of the general potential of the methods utilized.

## 4.2.2   State of the Method

In the past, studies frequently focused on the characterization of layers by applying an increasing number of new scientific methods, while the objectification and transparency of the original delineation of the layers was often neglected. One of the first systematic approaches to recording archaeological strata was introduced by Harris (1979), who invented the now widely applied *Harris Matrix*. The goal of Harris' approach is the creation of a sequence, an aim that has caused the Harris Matrix to be widely criticized (Renfrew and Bahn, 1996; Warburton, 2003). The Harris Matrix remains an important tool for stratigraphic documentation and most recent work – see, for example, the work of De Roo et al. (2016) – aims to achieve the automated implementation of the Harris Matrix into modern digital documentation techniques. Nevertheless, in most reference books the *act of layer identification* is mentioned but not further explained (e.g. Roskams, 2001). Despite taking the Harris Matrix as the first step in stratigraphic recording, most studies view the characterization of layers

as a highly complex process (e.g. Warburton, 2003), based on experience-based knowledge during macroscopic field analysis and supported by the application of quantitative methods such as analysis of the physical and chemical characteristics of the strata after concluding fieldwork. Currently lacking any quantitative measurement, the identification of strata during fieldwork remains a predominantly subjective act, highly dependent on the individual excavators and their experience and theoretical knowledge.

Most scientific assistive technologies that are used to characterize layers also help to compensate for this initial subjectivity (e.g., analysis of grain size). Likewise, they often involve great temporal, personal, and financial effort. Conversely, the application of semi-automated image analysis techniques potentially offers a fast and efficient way to support stratigraphic analysis during fieldwork. Based on physical characteristics captured via the reflectance signals, their statistical attributes offer in situ layer differentiation that is comprehensible and reproducible.

Hengl et al. (2007) and Lark (2005) argue that there is no automated system that can replace the *manual* identification of certain stratigraphic features (e.g., soil state), while they postulate the usage of consistent and unbiased methodology in combination with state-of-the-art tools, which are becoming more numerous with technological progress. Here, we therefore include the manual description and identification of strata, supported by quantitative geochemical analyses conducted by Böttcher (2002); this serves as reference data for the results of the quantitative analyses of image data.

Regarding spectral imaging, recent archaeological research mainly focuses on the application of airborne multispectral imaging for the mapping of archaeological crop marks and anthroposols (e.g. Aqdus et al., 2012; Retzlaff, 2015) or the detection of archaeological sites (cf. Comer and Harrower, 2013). While partly complemented by selective spectral measurements using handheld spectrometers, all these studies are concerned with a larger scale than a single excavation (e.g. Cheng and Wu, 2008; Schmid et al., 2008). Constant advancement of spectral imaging technology likewise improves the quality of recordings regarding spectral and spatial resolution, as well as the transportability of the systems, thereby extending the potential field of application.

Furthermore, spectral analysis is applied for color determination of soil and sediment samples (Ulery and Graham, 1993; Wills et al., 2007) and is also suitable for the detection of several mineral and elemental components (e.g. Jarmer and Schütt, 1998; Udelhoven and Schütt, 2000). In-situ analyses were recently conducted by Viscarra Rossel et al. (2009), who measured properties like soil color and clay content using a spectrophotometer. Hyperspectral imaging was recently proposed as a useful tool for extensive soil profile characterization (Steffens and Buddenbaum, 2012); however, these studies used hyperspectral scanners, and were thus limited to the analysis of sediment and soil data in the laboratory. Likewise, they lay the basis for future work: using extensive spectral imaging as an additional layer of documentation during fieldwork and thereby allowing for the analysis of the above-mentioned soil and sediment properties. At this point, the presented study is to be seen as a first step, evaluating the potential and applicability of the proposed methods.

**Figure 4.1:** Location of the sections examined (C). A: Cambisol with colluvial layers with artificial lighting; B: Luvisol in daylight conditions. Interpretation modified from Böttcher (2002); soil texture after Schoeneberger et al. (2012); C: Location of the soil profiles examined.
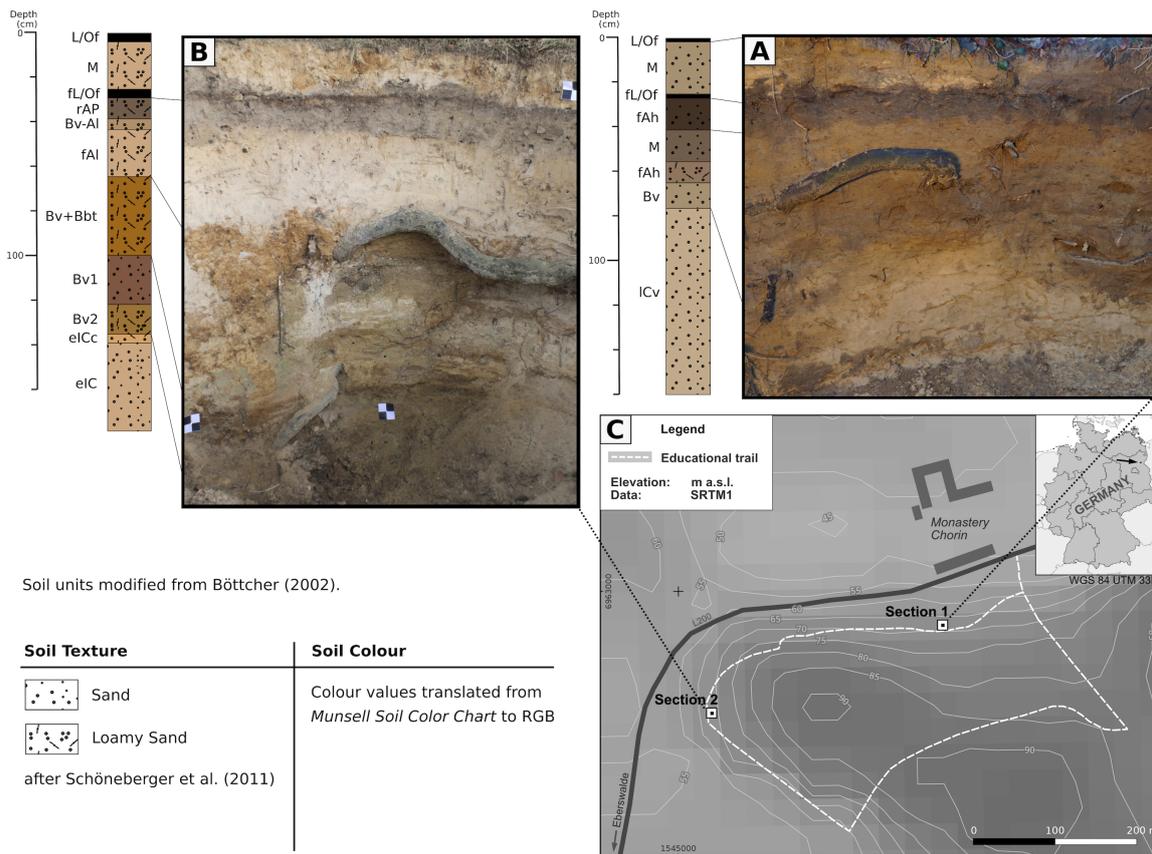
## 4.2.3   Material

### 4.2.3.1 Profile description

The Educational Soil Trail Chorin-Weinberg (Umweltbundesamt, 2010) is located in the area of the historic vineyard of the Cistercian Abbey Chorin (Brandenburg, Germany); two sites here were used as test profiles. The selected soil sections are a Cambisol with two colluvial layers (referred to below as Section 1) and a Luvisol (referred to below as Section 2)(Figure 4.1, Table 4.1). Description and interpretation of the profile data was adapted from Böttcher (2002), who characterized the strata in situ and additionally performed sedimentological analyses (e.g., chemical properties, pH-value, grain size). The delimited horizons are depicted in Figures 4.1 and 4.2. Both sections are situated in the landscape of the Angermünder Sander, which is why all horizons show a high sand content (> 70 %, Böttcher, 2002). Humus content of the topsoil is below 2.0 % in Section 2 and below 3.0 % in Section 1. The upper part of Section 2 (up to ca. 65 cm depth below surface, fAl, Figure 4.1) shows more acidic pH-values while the lower parts show pH-values greater than 6. Section 1 is completely acidic (pH < 5, Böttcher, 2002). The Cambisol (Section 1) is topographically situated at a footslope; as well as the two colluvial layers outlined by Böttcher (2002), it shows an additional sandy layer 25 cm thick that overlies the two colluvial layers. This topmost sandy layer most likely originates from modern anthropogenic reworking and is regarded as the result of maintenance works related to the

| | **Section 1** | **Section 2** |
|---|---|---|
| soil type | Cambisol (with two colluvial layers) | Luvisol |
| easting | 1545538.5 | 1545067.8 |
| northing | 6962931 | 6962750.3 |
| exposure | south-east | south-west |

**Table 4.1:** Characteristics and locations of the sections examined (CRS: WGS84 UTM 33N).

| | spatial resolution | spectral range | no. of bands | bandwidth | color depth |
|---|---|---|---|---|---|
| **Sony ILCE-6000** | 5696 x 4272 px | see Table 4.3 | 3 | see Table 4.3 | 8 bit |
| **Cubert X2 S258** | 512 x 272 px | 450(470) - 850(830) nm | 41 (37)* | 10 nm | 16 bit |

**Table 4.2:** Technical specifications of the camera systems used. *Four bands of the hyperspectral dataset were dropped due to high image noise.

soil trail. The same applies to the Luvisol in Section 2, where a 20 cm thick sandy layer lies on top of the Luvisol. In both sections, only a thin mull layer covers this anthropogenic layer (Figure 4.1).

### 4.2.3.2 Image aquisition

Digital RGB photographs of both sections were taken using a 24.3 MP mirrorless camera with a 2.8/16 mm lens (Sony ILCE-6000, Sony SEL-16F28). Hyperspectral imaging was carried out using a Cubert X2 S258 snapshot camera with a Schneider-Kreuznach Cinegon 1.8/16 mm lens. The two camera systems differ significantly in various technical details (Tables 4.2, 4.3), the most obvious being their spectral and spatial resolution. These two set-ups with different spectral and spatial resolutions were chosen deliberately to emphasize the challenges and opportunities of RGB and multispectral camera systems. While we compensated for different levels of lens distortion and image noise during image pre-processing (see Methods), we also took the differences as a chance to analyze how the different systems perform the task at hand. While the RGB camera has a far higher spatial resolution compared to the hyperspectral camera, the latter is able to capture more information in the spectral domain. The technical distinctions between the cameras used (Table 4.2) allow assessment of the impact of both spatial and spectral resolution on the results.

Due to the lack of information regarding spectral range and bandwidth of the Sony ILCE-6000 RGB camera, spectral properties of the camera were recorded by applying a systematic measurement of monochromatic light, emitted from a Newport MS257 monochromator using a halogen light source. The 8-bit proprietary raw data was normalized to the range [0, 1]. From the resulting data we were able to derive the central wavelengths of the red, green, and blue channels as well as the respective FWHM (full width at half maximum) and thus the approximate bandwidth (Table 4.3).

Special attention was given to the problem of uniform lighting conditions during fieldwork. Images were captured under natural daylight, as well as artificial light. Working with daylight was perceived as problematic for various reasons: parts of the image may be either over-illuminated or

| band | central wavelength | relative reflectance | FWHM |
|------|--------------------|----------------------|------|
| Red | 595 nm | 0.32 | 63 nm |
| Green | 540 nm | 0.42 | 93 nm |
| Blue | 485 nm | 0.16 | 78 nm |

**Table 4.3:** Central Wavelengths and FWHM (full width at half maximum) of the three image bands of the RGB camera set-up Sony ILCE-6000, Sony SEL-16F28. Relative Reflectance was measured using the 14-bit proprietary raw data and normalized to the range [0, 1]. The values shown are the results of a systematic measurement of monochromatic light.

shadowed; and depending on the time of the day and the on-site conditions, light color, and intensity may vary drastically between multiple recordings. To overcome these problems, the sections were recorded using a halogen light source (500 W). Depending on the size and lighting conditions of the section, either one or two lamps were used. Experimental usage of LED (light-emitting diode) lighting was carried out, but considered problematic and subsequently discarded since it does not deliver a continuous spectrum or a proper spectral signal above 700 nm, which is required for recording with the hyperspectral camera used here. The observed significance of artificial lighting corresponds with the experiment setup of Steffens and Buddenbaum (2013), for example, but adds the problem of in-field influences like irregular shadows contrary to image acquisition in the laboratory.

The imaging setup followed the scheme depicted in Figure 4.2. The cameras were placed on a tripod and centered in front of the respective section. Rotation of the cameras was limited to a minimum and black and white reference targets were placed in the soil sections to prevent spatial distortion and ensure comparability between the images recorded by the two cameras. Different angles of incidence were examined for the artificial light source. Depending on the local conditions, the final angle of incidence was chosen so that the casting of shadows was minimized. Hyperspectral images were recorded using the proprietary software Cubert Utils Touch, allowing for the export of the actual spectral reflectance in 16-bit TIF-format. The 8-bit RAW images of the Sony ILCE-6000 RGB camera were also converted to TIF-format. For further processing, the pixel values of both cameras were subsequently normalized to the range [0, 1] via feature-scaling and the respective color depth.

## 4.2.4  Methods

To ensure transparency, the data collected for this study are published together with reproducible scripts for all processing steps, written in the open-source programming language R (Online Resources 4.1 - 4.5).

### 4.2.4.1 Data pre-processing

To create overlaying datasets and thus to ensure comparability, hyperspectral and RGB images were rectified through the pairwise manual selection of reference targets and distinct geometrical features in the images, which were used as reference points for *thin plate spline* (TPS, Duchon, 1977) trans-
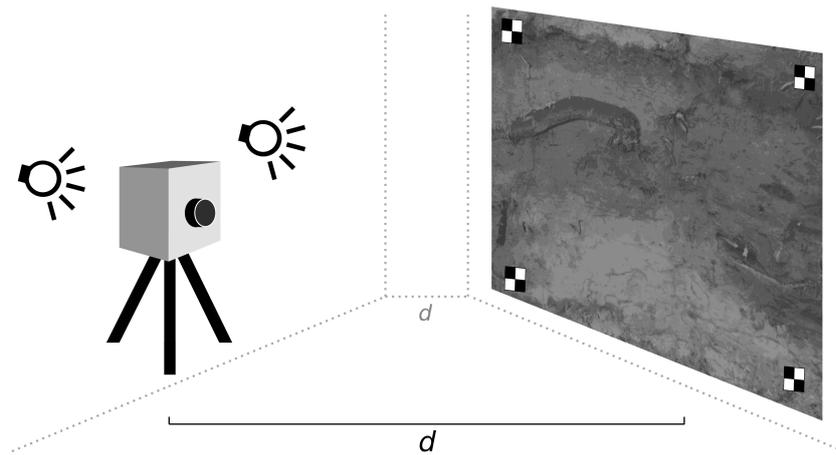
**Figure 4.2:** Schematic depiction of the imaging setup. Artificial light sources were placed near the camera position to gain uniform lighting conditions. The cameras were placed on a tripod, at a distance (d) of 1.5 to 2.5 m from the section. Reference points were placed at each section, allowing for overlapping image pairs.

formation with nearest neighbor resampling. These steps were carried out using the georeferencer-plugin for QGIS (v2.18). Rectification is mandatory to evaluate the quality of the results with respect to one set of lines, representing the initial visual stratigraphic interpretation.

Natural disturbances like big roots, stones, or inevitable shadows in the resulting image pairs were masked manually using the raster package (Hijmans, 2017) for the R programming language (v.3.4.3, R Core Team, 2017). Following Steffens et al. (2014), who applied Principle Component Analysis (PCA) to efficiently mask shadows in their hyperspectral recordings, the remaining vegetation was masked by defining a threshold value for one of the principle components, derived from an initial PCA of the raw data.

Examining the quality of the acquired hyperspectral images revealed that some bands had very high noise content, including 450 nm, 460 nm, 850 nm, and 840 nm. These bands were excluded from the analysis, resulting in a spectral range of 470-830 nm and a total of 37 spectral bands. Remaining spectral noise was minimized using a Savizky-Golay filter with a width of five spectral bands and a second-order polynomial (Savitzky and Golay 1964, following Lehnert et al. 2017).

Furthermore, both RGB and hyperspectral images were spatially filtered. The applied median filter was used to eliminate over-saturated pixels and likewise minimize the influence of an overly detailed capturing of the soil or sediment texture. However, as our primary aim was the delineation of the main strata, the potential loss of information is not noteworthy. Using the median rather than the mean ensured that the impact of outlying pixel values was minimized. Small stones and roots may otherwise have a significantly greater influence on the further processing of the high-resolution RGB images. Hyperspectral images were filtered using a 3 x 3 pixel moving window and RGB images using a 21 x 21 pixel moving window (window size was scaled to the respective spatial resolution; raster processing was conducted following Hijmans 2017). When dealing with cultural layers, characterized by a high content of materials such as charcoal or mortar, it may be more suitable to use a different spatial filter to prevent elimination of the spectral reflectance of these materials. However, this procedure is case-dependent and should be chosen according to the respective object
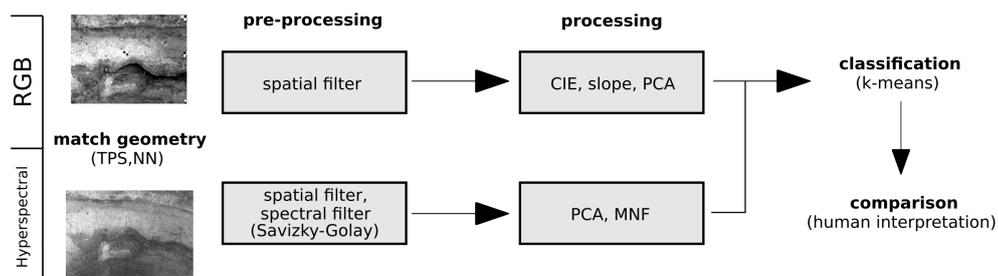
**Figure 4.3:** Schematic depiction of our workflow. Pre-processing included image rectification, masking of disturbances, spatial and spectral filtering. Different combinations of derivatives were tested: Principle Component Analysis (PCA), Minimum Noise Fraction (MNF), an alternative color space (CIExyY) and multiple spectral slopes. Subsequently unsupervised classification was carried out.

of interest: the spatial resolution of the hyperspectral camera may lead to mixed pixels, which would also influence the effect of the spatial filter (e.g. Hsieh et al., 2001).

### 4.2.4.2 Data processing

Multiple processing steps were tested to create derivatives of the raw data (Table 4.2, Figure 4.3). Taking the above-mentioned limitations of the available sets of data into account, we assessed the influence of the pre-processing and processing steps on the respective classification results.

**RGB data**   Processing of RGB data aimed to increase dimensionality. Additional rasters were created using PCA (conducted following Leutner and Horning 2017) as well as by calculating spectral slope values between the red, green, and blue channels. Two of the three principle components were included in the following analyses.

Attempts to compensate for uneven illumination via image processing were not successful. Therefore, conversion to the CIExyY color space (conducted following Gama 2016) and background illumination subtraction were dropped from the workflow.

**Hyperspectral data**   Several authors point out the significance of spectral resolution for the analysis of hyperspectral data (e.g. Dalponte et al., 2009; Peña et al., 2013). Due to the high spectral resolution (10 nm) within a comparatively narrow spectral range of the hyperspectral camera (470-830 nm) utilized, several processing steps were carried out in the spectral domain. These include Minimum Noise Fraction (MNF, applied following Bivand et al.2013 and Pebesma 2012) and PCA (applied following Leutner and Horning 2017). MNF and PCA were used to summarize and organize the spectral feature space, and to exclude redundant information. The first five principle components were used, since they explain 99.5 % of the variance of the data. The same number of MNF rasters was created, leading to a total of 10 raster bands derived from these processing steps (Table 4.4).

During image analysis a certain degree of autocorrelation between the created datasets became visible. Rather than being a problem however, this was considered useful since it helps to emphasize certain variations within the respective data (see Results).

| | no. of bands (RGB) | no. of bands (hyperspectral) | observed influence on classification results |
|---|---|---|---|
| **PCA** | 2 | 5 | more homogeneous clusters, in combination with raw data (RGB) or MNF (Hyperspectral) a better capturing of borders is visible |
| **MNF** | - | 5 | more homogeneous clusters, in combination with PCA (Hyperspectral) a better capturing of borders is visible |
| **slope** | 2 | - | less homogeneous clusters, helps to identify organic material in combination with raw data (RGB) and PCA (Hyperspectral) |
| **preprocessed data (PpD)** | 3 | 37 | especially spatial filtering has a high influence on cluster homogeneity |

**Table 4.4:** Number of raster bands created from processing steps and the respective observed influence on classification results. The best result was chosen manually with reference to the initial interpretation of the sections. For a visualization of the observed influence see supplemental material.

### 4.2.4.3 Image Analysis

The processed data were used for unsupervised classification, which does not require the definition of training areas and thus assumptions prior to the performance of the classification. We therefore propose unsupervised image classification as an objective addition to visual interpretation, avoiding the researcher's subjective perception as required for a supervised classification.

A simple approach of unsupervised classification is the k-means clustering algorithm, dividing the image data into $k$ different clusters based on either the complete data or a subset of it. In past studies this was successfully used for the analysis of soil data (e.g. Steffens et al., 2014). Clustering is carried out in the spectral domain and therefore does not take spatially adjacent pixels into account. Experimental usage of fuzzy c-means and artificial neural networks did not improve our results, which is why we propose the comparatively low computational effort of the k-means classifier as suitable for the purpose of this study. The k-means algorithm proposed by Hartigan and Wong (1979) was applied.

To analyze the impact of different pre-processing steps, band combinations, and derivatives, different feature sets were used as input for the image clustering (Table 4.5). As proposed by e.g. Peña et al. (2013), the pre-processed raw data is included in these combinations, since it may increase the quality of the results. Clustering was performed using a random sample, consisting of 10 % of the pixels of the respective raster stack (Hijmans, 2017; R Core Team, 2017). This reduction of the input data was carried out to reduce computational time. The allowed maximum iterations was set to 50 and three random sets were created as seed points. The resulting partition was used for prediction via the *clue* package (Hornik, 2005).

Experimental usage of different numbers of clusters revealed the best practice to be a total of 15 clusters. The number of clusters was chosen with respect to the number of manually identified stratigraphic layers: based on Steffens et al.'s (2014) work, a higher number of clusters was chosen than the number of layers, allowing for the grouping of clusters in the following step. The grouping of resulting clusters was performed manually and with respect to the visual stratigraphic interpretation.

| | |
|---|---|
| **Hyperspectral data:** | PpD; PpD/PCA; PpD/MNF; PpD/PCA/MNF; PCA/MNF; PCA; MNF |
| **RGB data:** | PpD; PpD/PCA; PpD/Slope; PpD/PCA/Slope; PCA/Slope; PCA; Slope |

**Table 4.5:** List of the tested combinations of bands, used as input for unsupervised classification. Abbreviations: pre-processed data (PpD), principle component analysis results (PCA), minimum noise fraction results (MNF) and spectral slope of the RGB data (slope).

While this step may be time-consuming and renders our approach *semi*-automated, it likewise holds several benefits: (i) grouping by the respective researcher can be considered a verification step in terms of his/her interpretation; (ii) indistinct borders of horizons or layers can be examined more precisely, resulting in a more solid argumentation as to *where* a border is drawn and *how* the border is shaped; and (iii) documentation of the grouping process allows for more transparent interpretation by the researcher.

### 4.2.4.4 Evaluation

Results were visually compared with the initial stratigraphic description (Böttcher 2002). Boundaries of stratigraphic layers as depicted in Figure 1 and Figure 4 are based on the horizons, which were defined by Böttcher (2002) via in situ characterization and sedimentological analyses. These layers were digitized as vectors (polygons) on the basis of RGB images of the sections. These vectors served as reference data for assessing the quality of the classification results. Evaluation was then carried out manually and by a visual interpretation of the homogeneity of the resulting clusters as well as their conformity with the delineation of stratigraphic layers provided by Böttcher (2002). Any absolute measurement of the deviation of the clustering findings from the initial interpretation would result in a semi-meaningful quantification, while the quality of the derived stratigraphic layers may vary significantly. We therefore propose that the manual comparison of the findings to the known stratigraphy is a better indicator for the quality of our classification results.

## 4.2.5   Results

The applied steps of image processing had different effects on the classification results (Table 4.4). Examining different combinations of the pre-processed data and the raw data, we found that the best results for RGB data were produced using a combination of the raw data, the PCA results, and the spectral slope rasters (Figure 4.4). The best results for the hyperspectral dataset were achieved using a combination of PCA results and MNF results (Figure 4.4). These combinations showed the most homogeneous clusters and captured stratigraphic delineations with high conformity to the visually based, manually drawn reference.

In artificial lighting conditions both camera systems capture the diversity of the examined sections adequately (Figure 4.4). Natural lighting leads to the confusion of unevenly lit areas with darker or lighter material. In general, both systems struggle to separate shadows from organic matter.

Additionally, the influence of pre-processing (spatial and spectral filters) on the results has to be taken into account. Experimental usage of different parameter sets showed that especially spatial filtering can significantly alter the quality of the results. Parameters were therefore chosen accordingly.
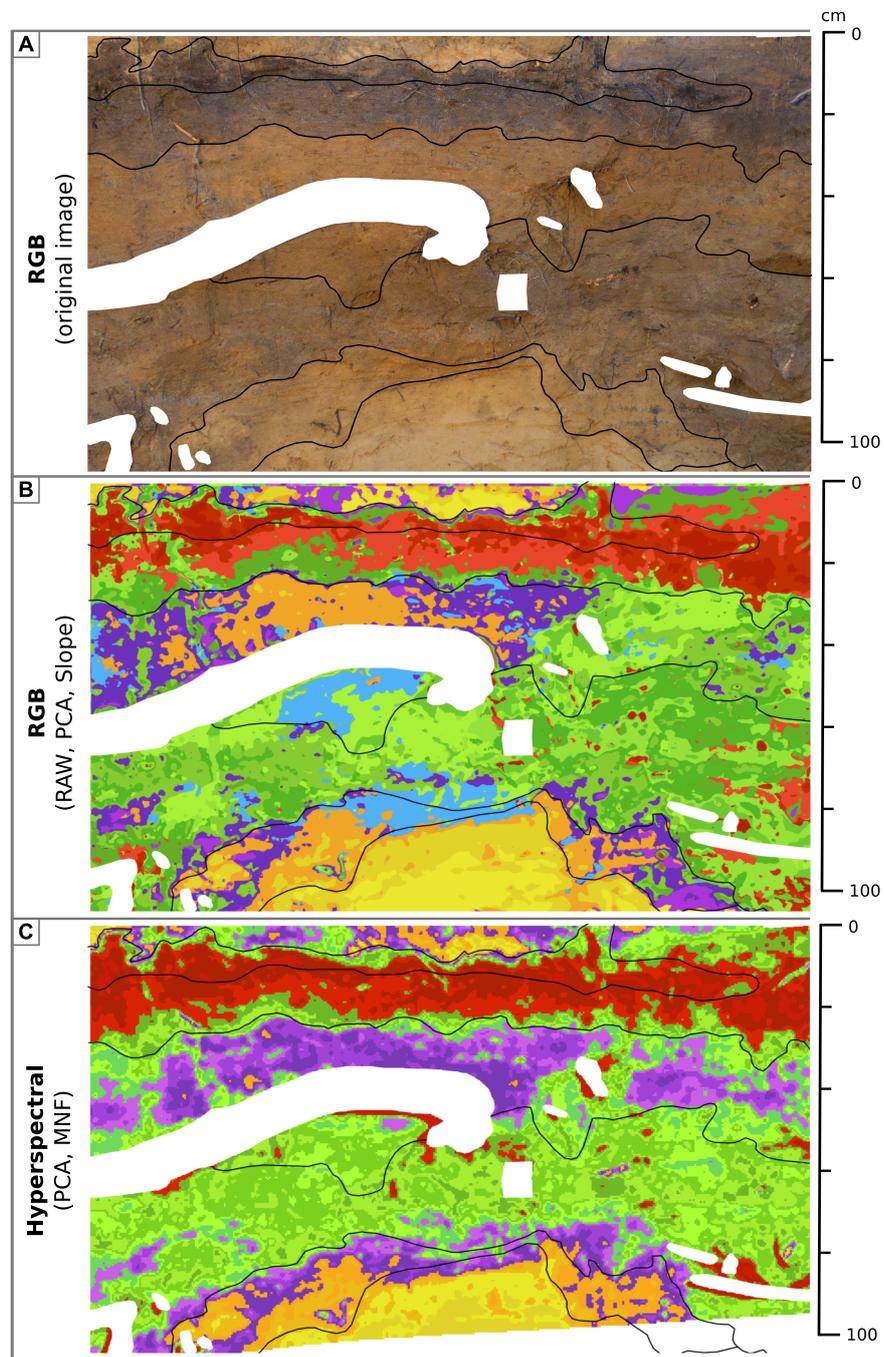
**Figure 4.4:** Manually grouped clusters resulting from unsupervised classification of image data captured from Section 1. Black lines indicate the initial visual interpretation. Colors indicate different cluster groups, individual stratigraphic features are shown in different hues of the group color. B: The proposed combination of pre-processed RGB data, Principle Components and spectral slopes. One cluster captured shadowed areas (black). C: The proposed combination of hyperspectral data: Principle Component Analysis (PCA) results and Minimum Noise Fraction (MNF) results. The hyperspectral image captured a smaller area than the RGB image, the lower part therefore remains transparent. Note the different performance of the hyperspectral data regarding slightly red material (highlighted in blue).

The differences between the two types of data at hand are very limited: both hyperspectral and RGB data are suitable to delimit the high-contrast strata via clustering with sufficient quality. While PCA and MNF derivatives of the hyperspectral data lead to a slightly better representation of organic material and allow for more homogeneous strata (Figure 4.4), the higher spatial resolution of the RGB data helps to identify smaller layers. Subsequently, our analysis shows that hyperspectral imaging in combination with artificial lighting has some benefits over RGB imaging. This result can be further emphasized since the hyperspectral data potentially holds additional information on the sediment properties. Nevertheless, our results indicate that for the sole goal of delineating strata in high-contrast stratigraphy, the comparatively low effort of processing RGB data can deliver results of similar quality and explanatory power.

Furthermore, the different performances of the methodology regarding the two case-study sections need to be examined. While hyperspectral imaging in combination with artificial lighting produced good results for Section 1 (Figure 4.4), it revealed significant problems in Section 2 (Figure 4.5). The classification results of the hyperspectral data reveal a misinterpretation of the slightly red material in Section 2 (Bv+Bbt horizons, Figure 4.1), clustering it together with the far brighter material of the fAl horizon, for example. This issue also appears in the clustered RGB data of the same section (Figure 4.5). Summing up, none of the applied imaging and processing steps were able to compensate for this problem and allow reliable delineation of Bv and Bbt horizons. Furthermore, the comparatively poor spatial resolution of the hyperspectral camera leads to small features being neglected (e.g., Section 2, fL/Of horizons, Figure 4.1). Processing of the data did not lead to a significant improvement.
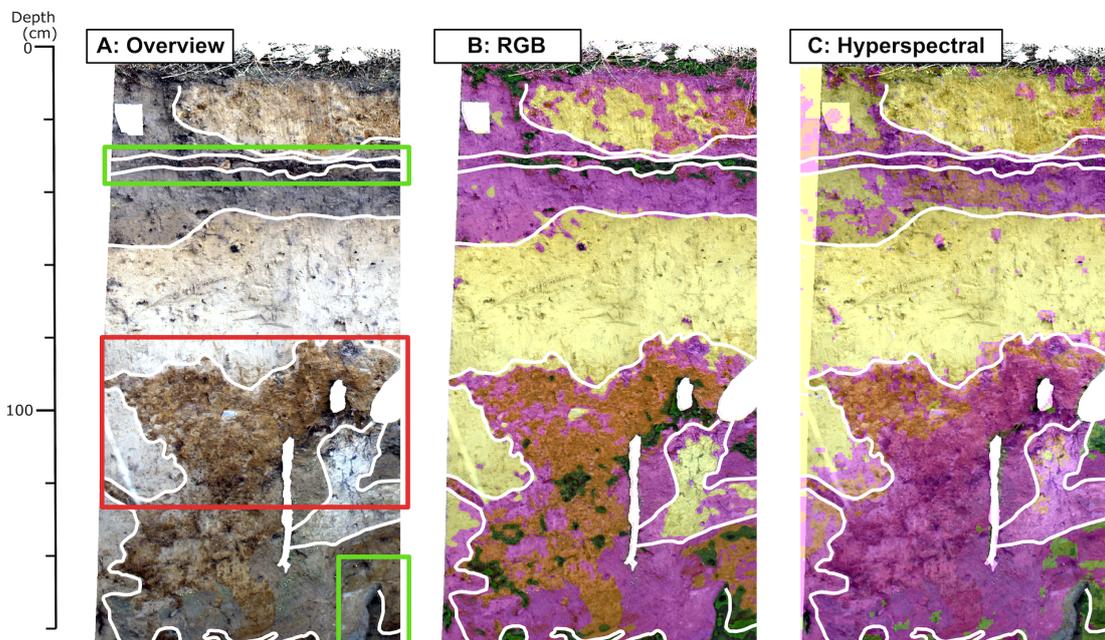


**Figure 4.5:** Grouped clusters resulting from unsupervised classification of image data captured from Section 2 (B: spatially filtered RGB data; C: composite of filtered hyperspectral data and MNF results). White lines indicate the initial visual interpretation. Colors are used to represent the cluster groups. The comparison highlights the problem of delimiting red material in the hyperspectral domain (red box) and the general mixing up of organic matter and shadowed areas (green boxes). Images were captured with natural lighting conditions.

## 4.2.6 Discussion

Our results indicate that relevant clusters can be derived from both standard RGB images and hyperspectral images; for the sole goal of delimiting stratigraphic features, RGB imaging delivered similar results to the hyperspectral imaging setup. These results are contradictory to our expectations since a finer spectral resolution is offered by the hyperspectral camera utilized. The PCA of the hyperspectral data showed that 99.5 % of the variance of the data can be explained by the first five principle components. Contrary to our results, these five components should still contain more information than the three channels of the RGB camera and therefore lead to better results when clustering the data (e.g. Dalponte et al., 2009; Peña et al., 2013). Likewise, the PCA results imply high noise content in the hyperspectral data, which in combination with the comparatively low differences between the classification results of the two datasets suggests that an in-depth review of the hyperspectral sensor is necessary. Other possible reasons for the disappointing performance of the hyperspectral camera include its lower spatial resolution, which may lead to mixed pixels (e.g. Hsieh et al., 2001), therefore complicating the visibility of spectral features. Additionally, large differences in the brightness of strata may interfere with the spectral signal of the examined material. Clustering strata of high-contrast soil and sediment profiles therefore heavily relies on differences in the brightness/albedo of the strata (e.g. Rees, 1990), which may explain the good performance of the RGB data. Furthermore, the wavelength range of the hyperspectral camera utilized (470-830 nm) only captures a few significant spectral features present in soil and sediment material. Apart from, for example, several iron-related spectral features between 500 and 640 nm, several spectral features beyond the range of the camera can be expected to be helpful for the analysis of soil and sediment profiles in future studies (e.g., iron absorption bands at 870 nm and around 1000 nm or spectral features of clay minerals beginning at around 1700 nm (Baumgardner et al., 1985)). This, however, would require the usage of a hyperspectral scanner, necessitating a far more complicated workflow during fieldwork.

For the RGB data, a relatively low weighting of the blue part and a high weighting of the green part of the spectrum is evident (Table 4.3). However, this only leads to small differences to the image analysis results gained using the hyperspectral camera. Capturing RGB images requires proper artificial lighting conditions since the results vary drastically between natural and halogen light. While natural daylight potentially delivers the best spectral reflectance, since it covers the full electromagnetic spectrum (e.g. Strahler, 1975), the vertical alignment of sections as well as recurring changes in the intensity and color of the daylight render a standardized artificial light source indispensable. We were not able to compensate for the disadvantages of natural daylight by means of data processing. This limits the applicability of the proposed method to sites where homogeneous natural lighting is present or an adequate power supply is available for operating an artificial light source.

LED lighting was dropped from this study due to its bad spectral signal in the hyperspectral images. Recent work by e.g. Islam et al. (2017) or Belušič et al. (2016) has examined how a combination of multiple LED lamps could be used for spectral imaging. However, these experimental constructions are not yet applicable under fieldwork conditions. Nevertheless, the good performance of RGB clustering may favor the use of an energy efficient LED light source since the spectral signal could be sufficient for RGB imaging.

The misinterpretation of shadows and organic matter, as present in Section 2, shows the pos-
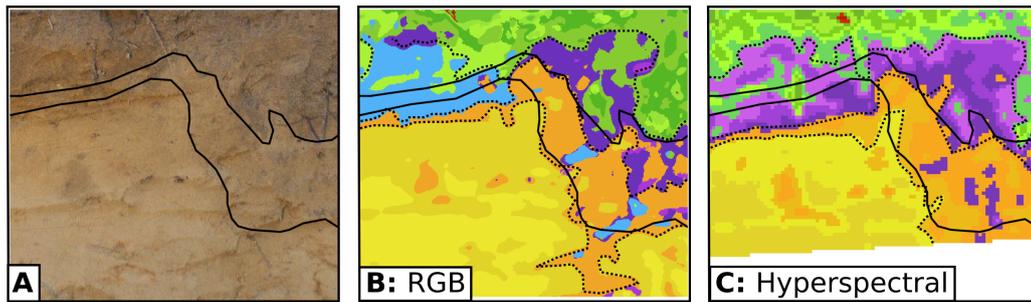
**Figure 4.6:** Comparison of stratigraphic interpretation (lower part of Section 1). White lines represent the initial visual interpretation, black dotted lines are manually added borders based on the classification results. A: in-field visual perception; B: RGB classification results; C: hyperspectral classification results. See Figure 4.5 for further description.

sible impact of different lighting conditions on the output of the image analysis. Given laboratory conditions, Steffens et al. (2014) were able to remove shadowed areas via PCA. Our results gained under field conditions show that during fieldwork these shadowed areas can prove highly problematic, which is consistent with the general problem of uniform lighting in spectroscopy (e.g. Rees, 1990). As a workaround, strongly shadowed areas may be masked manually prior to the analyses. Compensation for uneven lighting via image processing should be handled with care and is best limited to the affected areas since it may otherwise alter the spectral signal of the whole image (e.g. Shepherd and Dymond, 2003).

Additionally, the distance between camera and section as well as the viewing angle should be chosen accurately to allow for a true-to-scale rectification of the digital images (Bryan et al., 2003). Reference points with known coordinates have proven suitable to compensate for possible image distortions, which is consistent with common practice during archaeological fieldwork (e.g. Sapirstein and Murray, 2017; Bernatchez and Marean, 2011). The quality of the results strongly depends on the imaging setup. We propose the setup depicted in Figure 4.2 is used to minimize the casting of shadows and image distortion. As recent advancements in the field of structure-from-motion indicate, these factors could be even better compensated for by rendering an orthophoto (e.g. De Reu et al., 2014).

Our results imply that hyperspectral imaging delivers no clear benefit for the identification of stratigraphic features and therefore are not in agreement with the results of e.g. Sorenson et al. (2018) or Steffens and Buddenbaum (2013), who were able to spatially map distinct soil properties via the spectral signal. Comparability with these results, however, is limited since the aforementioned studies took place in the laboratory and were concerned with supervised classifications of soil properties rather than the unsupervised delineation of strata. Additionally the spectral camera used in our study covers a smaller spectral range than the one applied in the studies of Sorenson et al. (2018) or Steffens and Buddenbaum (2013). At this point, an in-depth review of this particular spectral camera would be necessary. As this would include a review of the potential information which can be captured in the spectral domain (e.g., iron oxide content (Baumgardner et al., 1985), soil moisture (Haijun et al., 2017), or organic carbon content (Steffens et al., 2014)), this topic is beyond the scope of the paper at hand.

Within this study, the processing of the data could not help to overcome some of the discussed disadvantages of the hyperspectral camera used. Various combinations of pre-processing and process-

ing steps revealed that the applied pre-processing methods have a far higher impact on classification results than the processing of the data via PCA or MNF. These results emphasize the high influence of spatial and spectral filtering on the data as is pointed out by e.g. Hsieh et al. (2001) or Vaiphasa (2006). We therefore agree with past studies and suggest the parameters of spatial and spectral filtering be selected attentively and with respect to the spatial and spectral resolution of the respective dataset.

### 4.2.7 Conclusions

We propose that the results of this study indicate the benefits that semi-automated image analysis can offer for stratigraphic interpretation. The methods used can help to render interpretation more comprehensible and reproducible, and thereby strengthen objectivity. The applied analysis of image data should be seen as a tool, emerging from technological advance and containing a potential gain of knowledge in the sense of Spinner (1974). At the same time, existing conventions should not be discarded. We rather suggest the proposed methodology as a possibility to verify and support existing ways of delimiting diagnostic horizons during fieldwork, which are based on cognitive-and thus subjective and irrational-individual work. A critically rational attitude, as requested by Spinner (1974), in combination with the proposed methodology permits increased accuracy and certainty when identifying diagnostic horizons by overcoming disciplinary borders (cf. Feyerabend, 1983). According to Morgan and Wright (2018), the documentation of archaeological features by drawing is at a "critical juncture" (p.145), between digital and analogue recording. In this context we see our study as a step towards an integrated approach which benefits from both analogue and digital techniques by offering digital assistance for the human eye.

Spectral imaging and parameterization complemented by statistical analysis allowed a delineation of clusters based on physical properties in the examined soil profiles. These clusters cannot be directly taken as stratigraphic layers. Rather, the clusters result from mathematical processing of physical measurements and offer specific features. Hereby an assistance of the human observation is provided which supports the recording of profiles and makes perceptive decisions of researchers more transparent. The clusters consisting of similar pixels thereby clearly help to assess diversity within the examined sections. We propose that the term *stratigraphic features* be used for these clusters: a subunit of stratigraphic layers which helps to draw and characterize borders and the homogeneity of layers. Added to the digital documentation process by using software such as CAD (computer-aided design), these features help to render interpretations more transparent. However, the quality of the clustering is assessed by a direct comparison with profile delineations based on expert knowledge; we have not yet regarded the possible additional information resulting from a clustering with higher differentiation than the expert-knowledge-based layer delineation.

By interactively grouping the resulting clusters, the researcher is offered the possibility to critically reflect his/her initial interpretation. This likewise allows for a better understanding of the stratification (Figure 4.6). On the other hand, the proposed workflow still requires a lot of manual work and therefore is open to discussion. Further automation could render the proposed method more

suitable for fieldwork. Likewise, we encourage everyone who makes use of this method to document the manual grouping of clusters to ensure reproducibility, a property that interpretation should always be concerned with.

The applicability of this method for archaeological and sedimentological fieldwork is yet to be determined. Several drawbacks (e.g., problems with even lighting, limited spectral range) have to be considered but may be solved by future technological advancement and further standardization of the method. The publication of our processing workflow in the form of reproducible scripts should nevertheless facilitate its adaptation to archaeological or sedimentological sets of data (see supplemental material). Anthropogenic layers are often characterized by a large amount of small debris of charcoal, bones, brick or mortar as well as a high content of organic material. Considering these characteristics, and depending on the excavation concerned, a higher weighting of sediment texture may be useful to distinguish between layers during archaeological fieldwork. Our experimental attempt regarding this topic comprised the analysis of Grey-Level-Co-occurrence-Matrices (GLCM, Haralick et al., 1973; Zvoleff, 2016), which led to higher image noise, therefore aggravating the delimiting of layers in our study. However, the comparably low variance of the overall sandy material of the soil sections may explain the poor results regarding image texture. For the analysis of the high-contrast soil profiles the parameter image texture was therefore not implemented in our workflow.

The discussed results highlight that properly recorded RGB images can be sufficient for the identification of stratigraphic features. A comparatively low amount of pre-processing produces reliable results, which are easy to achieve with regular fieldwork equipment. Digital photography is nowadays a natural part of the standard workflow during excavations. In combination with our script-based approach, this implies that results similar to ours can be achieved with comparatively low additional effort. The results also connote that for examining high-contrast stratigraphy, in an extreme case, a high-quality grayscale image could be sufficient to delimit the strata via unsupervised classification. Our approach should therefore be seen as a motivation for using the acquired photographs as physical measurements that hold potential information. The proposed method should be considered an addition to the established methodological spectrum of archaeological and sedimentological documentation and interpretation. We were able to develop the basis for a tool which allows the black box of individual subjective perception during fieldwork to be partially overcome. Additionally, the open-source character of the software allows a fast validation of the applicability of the method within different scenarios, like archaeological excavations.

While the identification of *stratigraphic features* for the assistance of human vision can be achieved with easy-to-apply RGB imaging, we want to stress the significance that hyperspectral imaging may have for the characterization of these features. As discussed above, past studies revealed multiple connections between a material's spectral properties and its physical and chemical composition. The snapshot character of the hyperspectral camera used, as well as the ongoing advancement of hyperspectral imaging technology, may therefore be utilized to map features of the soil matrix, such as mineral concentrations, within sections and plans in a spatially explicit way. Further studies regarding this topic are in preparation.

Finally, we want to point out that the quality of the stratigraphies derived through statistical analysis of RGB and hyperspectral imagery is proven by the comparison to the delineations received

by traditional field and laboratory survey; this, in general, means that, based on the on-site layer delineation, samples are extracted and are analyzed in the laboratory. It is common practice that these samples are extracted in high density and that the final layer delineation is accomplished based on the physical and chemical sediment characters. Thus, conducting quality control as described, we assume the on-site delineation of layers by a scientist to be "correct" and likewise exclude the possibility that other non-visible parameters may lead to a different output regarding the present stratigraphy.

## Acknowledgments

CHAPTER 5

# VIS-NIR spectroscopy and laboratory sediment analyses

## 5.1  Introduction

Established methods for delineating stratigraphic layers quantitatively often require time, money and expert knowledge. As shown in chapter 4, the statistical analysis of digital image data constitutes a promising alternative to be used for this purpose. However, there remain several issues which need to be addressed: (i) how does image analysis perform, when compared to laboratory sedimentological analyses? (ii) Are there situations where one of these approaches should be considered more useful than the other? (iii) How does the workflow proposed in chapter 4 perform when applied to a more complex archaeological profile? And (iv) can this workflow be optimised by additional processing steps?

The following study tries to answer these questions by comparing traditional laboratory sediment analyses with the results of a quantitative analysis of selective and extensive spectral measurements. An archaeological profile was recorded with a RGB and a multispectral camera. Additionally a high number of sediment samples was extracted from the profile and analysed in the laboratory. Besides common sedimentological analyses, we additionally measured the spectral reflectance (350 - 2,500 nm) of each sample. Statistical analyses of these datasets show a promising performance of both the selective and the extensive spectral data when compared to often destructive and expensive laboratory analyses.

## 5.2  Paper 2: On-Site VIS-NIR Spectral Reflectance and Colour Measurements – A Fast and Inexpensive Alternative for Delineating Sediment Layers Quantitatively? A Case Study from a Monumental Bronze Age Burial Mound (Seddin, Germany)

### Abstract

Quantitative sediment analyses performed in the laboratory are often used throughout archaeological excavations to critically reflect on-site stratigraphic delineation. Established methods are, however, often time-consuming and expensive. Recent studies suggest that systematic image analysis can objectivise the delineation of stratigraphic layers based on fast quantitative spectral measurements. The presented study examines how these assumptions prevail when compared to modern techniques of sediment analysis. We examine an archaeological cross-section at a Bronze Age burial mound near Seddin (administrative district Prignitz, Brandenburg, Germany), consisting of several layers of construction-related material. Using detailed on-site descriptions supported by quantitatively measured sediment properties as a measure of quality, we compare clustering results of (i) extensive colour measurements conducted with an RGB and a multispectral camera during fieldwork, as well as (ii) selectively sampled sedimentological data and (iii) visible and near infrared (VIS-NIR) hyperspectral data, both acquired in the laboratory. Furthermore, the influence of colour transformation to the CIELAB colour space (Commission Internationale de l'Eclairage) and the possibilities of predicting soil organic carbon (SOC) based on image data are examined. Our results indicate that quantitative spectral measurements, while still experimental, can be used to delineate stratigraphic layers in a similar manner to traditional sedimentological data. The proposed processing steps further improved our results. Quantitative colour measurements should therefore be included in the current workflow of archaeological excavations.

**Keywords:** digital archaeology; image classification; stratigraphy; landscape archaeology; sediment analysis; spectroscopy; proximal sensing

### 5.2.1  Introduction

Stratigraphic documentation and interpretation are crucial parts of archaeological research. Since the emergence of the archaeological discipline, researchers have been concerned with the identification of methods that allow more objectivity in stratigraphic delineation and interpretation. The fundamental work of, e.g., Harris (1979) is meanwhile complemented by a progressing digitisation of fieldwork (e.g. Vincent et al., 2014) and a growing number of on-site measurements and analyses carried out in various laboratories in the aftermath of an excavation. These techniques comprise, e.g., the establishment of soil colour standards (cf. Simonson, 1993; Landa and Fairchild, 2005), geochemical analyses of sediments (e.g. Hassan, 1978; Goldberg et al., 2013; Hoelzmann et al., 2017), the analysis of thin

sections (e.g. French, 2003), or the analysis of pollen (Kooistra and Kooistra, 2003) and phytoliths (Meister et al., 2017).

The overall objectivity of stratigraphic interpretation clearly benefits from these techniques. Nevertheless, the initial on-site delineation of stratigraphic layers (i.e., stratigraphic documentation) remains partly influenced by the perception of the respective researcher at work. As pointed out by Zhang and Hartemink (2019b) and Haburaj et al. (2019), statistical analysis of RGB and multispectral images of soil profiles can be utilised to semi-automatically delineate stratigraphic units. These derived units are based on the physical measurement of the sediments' spectral reflectance and allow the respective researcher to critically face up to his or her delineation.

Building upon these developments, our study further explores the potential of visible and near-infrared (VIS-NIR, 400–700–2500 nm) spectral data for archaeological fieldwork. Due to its non-destructive character and the moderate extent of necessary sample preparation, spectroscopy in general has the potential to save time and costs during fieldwork (e.g. Stenberg and Viscarra Rossel, 2010; Debret et al., 2011; Zeeden et al., 2017; Haburaj et al., 2019). Working with an archaeological cross-section from a Bronze Age burial mound from Brandenburg, northern Germany, we investigate how traditional geochemical sediment analyses, selective spectroscopic measurements, and extensive RGB and multispectral image data can be used to objectivise archaeological field documentation by quantitatively capturing stratigraphic layers. Using a detailed on-site description supported by quantitatively measured sediment properties (grain size composition, water content, and soil organic carbon content) as a measure of quality, we compare clustering results of (i) extensive colour measurements (i.e., image data) acquired with an RGB and a multispectral camera during fieldwork, as well as clustering results of selectively sampled (ii) portable energy-dispersive X-ray fluorescence spectrometer (p-ED-XRF) data and (iii) visible and near infrared (VIS-NIR) hyperspectral data, both acquired in the laboratory. This experimental approach allows us to illustrate the benefits and shortcomings of systematic spectral measurements as a complementary or alternative method for stratigraphic documentation, delineation, and interpretation.

The characteristics of the recorded diffuse reflectance spectra of the sediments are directly linked to their chemical and physical composition (e.g. Stenberg and Viscarra Rossel, 2010). Within the last two decades, an increasing number of publications were devoted to the prediction of soil properties via VIS-NIR spectroscopy (e.g. Udelhoven et al., 2003; Mutuo et al., 2006; Brown et al., 2006; Sankey et al., 2008; Volkan Bilgili et al., 2010). Soriano-Disla et al. (2014) deliver an extensive overview on the predictability of multiple soil properties via visible, near-infrared, and mid-infrared spectral data. We therefore examined how the conducted image classification performs when using additional input data: We were able to calibrate a prediction model for soil organic carbon (SOC) based on the recorded spectral signal and created a raster containing predicted SOC values for the entire profile. Several authors have argued that the CIELAB colour space (Commission Internationale de l'Eclairage) shows a clear benefit regarding quantitative analyses (e.g. Melville and Atkinson, 1985; Viscarra Rossel et al., 2006; Jarmer et al., 2009; Hill et al., 2010; Sprafke, 2016; Vodyanitskii and Kirillova, 2016). The CIELAB colour space consists of three values: L* is the vertical axis and is defined by lightness, while a* and b* represent the chromaticity (a*: Red/green; b*: Yellow/blue). We examined colour transformation to the CIELAB colour space, as it was found to be highly effec-

tive when analysing soil profiles (e.g. Zhang and Hartemink, 2019a). Hereby, our overarching aim remains the delineation of stratigraphic layers rather than the quantitative examination of sediment properties.

Our results indicate the importance of quantitatively measuring sediment colours—be it selective or extensive—as they constitute an additional layer of documentation during an archaeological excavation, which can be acquired easily and render the on-site archaeological documentation and interpretation more transparent and reproducible. The study at hand is presented as an experimental investigation, which allows us to assess the potential of VIS-NIR spectral sensors for archaeological fieldwork.

## Study Site

The Bronze Age burial mound of Seddin (colloquial: 'King's Grave', German: Königsgrab Seddin), which was discovered in 1899, is widely known for its tremendous size and the valuable funerary goods made from bronze, iron, glass, and stone (e.g. Kiekebusch, 1928; May, 2003; May and Hauptmann, 2012). It was dated to approximately 800 BC (period V of the Nordic Bronze Age) and has a central role in the regional cultural landscape, which today covers the north-west of the federal state of Brandenburg and southern Mecklenburg (Wüstemann, 1974; May, 2018b). While past studies found the surroundings of the burial mound to be heavily shaped by human influence, which occurred during the Bronze Age (Schenk and Goldmann, 2004; May and Hauptmann, 2012), the 'King's Grave' keeps an outstanding position, both regarding its geographical location and its rich physical characteristics. Estimates by May (2018a) see the total height of the mound between eight and ten metres. The burial mound is separated spatially from smaller groups of burial mounds in the surrounding area. It is located on a slight spur, surrounded by multiple waterbodies (May and Hauptmann, 2012): 900 m east of the burial mound runs the river Stepenitz, and in the north and south, two of the river's tributaries delimit the burial mound from its surroundings (Figure 5.1). The parent material is dominated by Pleistocene deposits, composed mostly of fine to coarse glaciofluvial sand with little or no gravel (Landesamt für Bergbau, 1905). West and south of the burial mound, clayey to sandy silt can be found. In the hollows, associated with the two tributaries of the Stepenitz (Figure 5.1), muck soil of a sand–humus mixture dominates, which overlies fluvial and periglacial deposits (Landesamt für Bergbau, 1905). The relief of the surrounding landscape is characterised by a smooth, rather levelled topography that is dissected by depressions and river valleys. Regarding the wider vicinity, several burial mounds of similar size are known but have mostly been destroyed and lack any archaeological record (Götze, 1912; May and Hauptmann, 2012). Regardless of the nature of these undocumented burial mounds, their sheer amount reflects the importance of this landscape during the late Bronze Age. Brunke et al. (Brunke et al., 2016) see the 'King's Grave' of Seddin as the clearest manifestation of an accumulation of wealth and power in this area around the middle reaches of the Stepenitz.

Within this study, we examine section SD17P1 (297601 E, 5891584 N; WGS84; UTM 33N), which was documented in 2017 during an excavation on the north-western slope of the Bronze Age burial mound of Seddin. The investigated section faces north, is approximately 3.2 m high, and consists mainly of anthropogenic layers: Only the lowermost 20–40 cm are interpreted as the buried
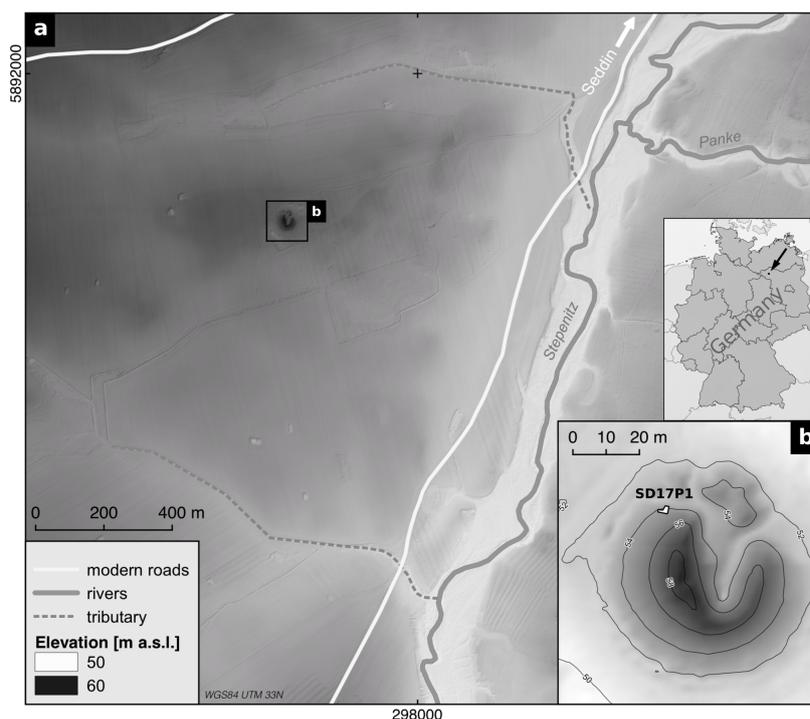
**Figure 5.1:** Location of the study site: (**a**) Surroundings of the burial mound; (**b**) location of section SD17P1. Elevation data: LiDAR-based DEM (1 × 1 m) by Geobasis-DE/LGB 2013 (R. Hesse; modified).

local soil (Figure 5.2c,E), which was dated by three ^{14}C-AMS ages to the ninth century BC (May, 2018b). The remaining part of the profile consists of discrete layers, which were piled up during the construction of the burial mound with alternating layers of stone and sand that were deposited on top of each other. In addition, there is a horizontally running, thin, dark layer, which crosses the lowest artificial sand layer (Figure 5.2c,C). The type of construction with stone layers at hand is congruent with the findings of Brunke et al. (2016), who examined the south-eastern side of the burial mound in 2013. However, the diagonal segmentation and orientation of sand layers B and D is only present in SD17P1. The topmost stone layer of SD17P1 most likely represents the Bronze Age surface, which would, in addition to the prominent position of the mound, emphasise its monumental character (cf. May, 2018b).

All analyses carried out in this study focus on the parts of the profile that lie above the lowest layer of stones, which is approximately 200 cm below the surface of the burial mound. Layer E (local soil) is therefore excluded from the processing (Figure 5.2c).

## 5.2.2   Materials and Methods

Our experimental design involves the measurement of selective data based on sediment samples extracted from the archaeological profile and extensive image data acquired during fieldwork. The workflow is depicted in Figure 5.3, and a more detailed documentation of the data processing along with scripts for the programming language R is given in Online Resources 5.1 - 5.3 (in Appendix).
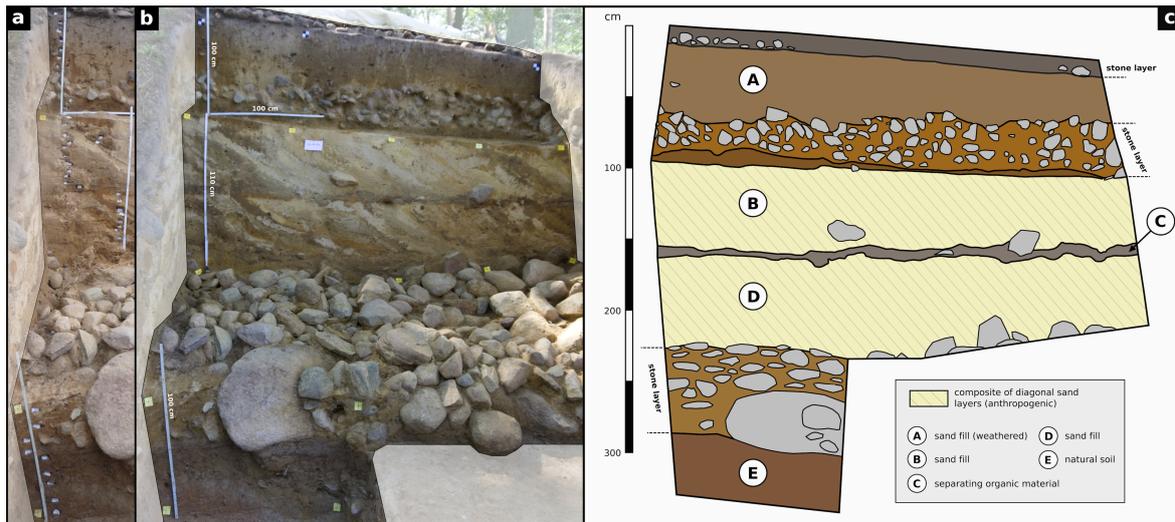
**Figure 5.2:** Archaeological section examined. (**a**) Location of the sediment samples taken; (**b**) overview of the section; (**c**) drawing of the archaeological section examined. Layers A to D are part of the an- thropogenically created burial mound. Layers B and D consist of numerous diagonal layers, which are summarised here. Colour and texture of the sediments were documented during fieldwork and are pre- sented in Figure 5.5 along with the sedimentological record (Section 5.2.3). Layer E is interpreted as the local soil that was buried during the construction of the burial mound (May, 2018b) and is excluded from all analyses.

## 5.2.2.1 Laboratory Sediment Analyses

Analyses were performed using the topmost 31 samples, extracted from the eastern part of the profile and vertically covering a depth of 6 to 194 cm, thus excluding the natural soil of layer E (Figure 5.2a). Samples were extracted and stored in airtight bags to minimise errors when measuring water content. We densely sampled layers B and D to consequently capture the variability between the single obliquely bedded sand layers. Stratigraphic (sub)layers were described regarding colour and texture using the Munsell soil colour chart (Munsell Color Company, 1994) and the German manual of soil mapping (Ad-hoc Arbeitsgruppe, 2005) respectively.

Further laboratory analyses of the sediments followed after the determination of the water content and sample preparation, including crushing of aggregates, separation of coarse components with a 2- mm sieve, and homogenising of the subsamples for carbon and portable energy-dispersive X-ray fluorescence spectrometer (p-ED-XRF) analyses in a vibrating disc mill.

**Water Content**    The water content was determined gravimetrically. The moist samples were weighed, heated in a drying cabinet at 105 °C until a constant weight was reached, cooled down to room temperature in a desiccator, and weighed again. The water content was calculated according to Blume et al. (2011) and reported in mass%.

**Soil Organic Carbon**    The total carbon content of the samples was determined using a TruSpec CHN analyser by Leco (St. Joseph, Michigan, USA). Additionally, a small portion of the powdered material from all samples was tested for its carbonate content using 9.9% HCl acid. None of the samples showed signs of carbonate, which is why the measured total carbon content is interpreted as
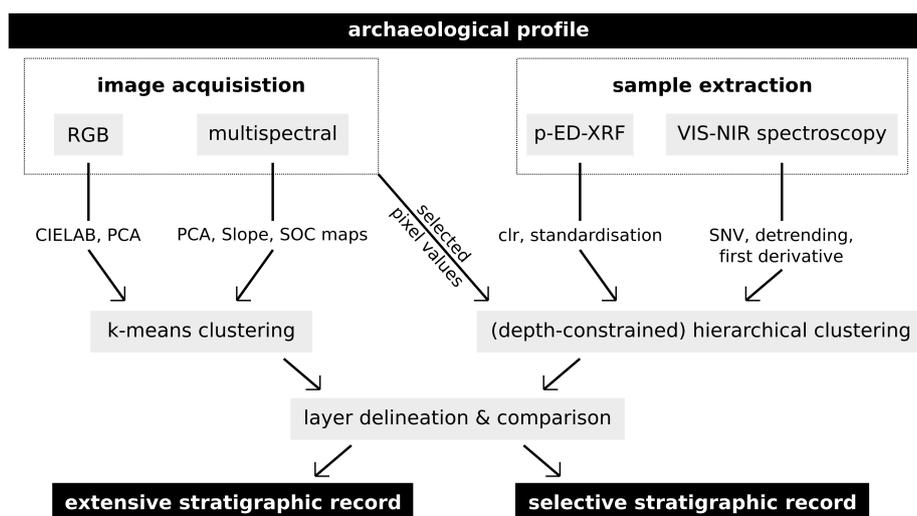
**Figure 5.3:** Schematic depiction of our workflow. Image processing included principal component analysis (PCA), derivation of slope rasters, colour transformation to CIELAB (Commission Internationale de l'Eclairage), and the prediction of soil organic carbon (SOC). The standardised centred log ratio (clr) of the data acquired with the portable energy-dispersive X-ray fluorescence spectrometer (p-ED-XRF) was used. Spectroscopic measurements of the sampled material were processed using standard normal variate (SNV) transformation with detrending applied to the first derivative of the $log(1/R)$ transformed spectrum. This allowed us to create stratigraphic delineations based on extensive and selective data and to compare the performance of colour and spectral measurements to traditional p-ED-XRF data.

the total organic carbon content. During measurement, the samples were dry combusted at 950 °C in an oxygen atmosphere, and carbon fluxes were quantified by infrared spectroscopy. The results were calibrated using certified reference material (CRM; Leco 502–309; 12.29 ± 0.37 mass% C; Leco 502–308; 3.6 ± 0.29 mass% C; reproducibility was <2% RSD) and are reported as soil organic carbon (SOC) in mass%.

**pH Value** The pH values of the samples were measured in a 1:2.5 solution of 10 g of dried sediment and 25 mL of 0.01 M KCl using a handheld pH meter (Hanna Instruments, Woonsocket, Rhode Island, USA) with a resolution of 0.1.

**Grain Size** The particle size distribution for the fraction $\leq 1$ mm was determined with a laser diffraction particle size analyser LS13 320 by Beckman-Coulter (Brea, California, USA). The samples were sieved with a 1-mm sieve and bi-distilled water together with about 0.5 g $Na_4P_2O_7$ as an anti-coagulation agent. The samples were placed in an overhead shaker for 24 h and ultrasonic treated for 5 min directly before the measurement. The prepared samples were put into a liquid sample divider and two subsamples were measured with three independent runs each. The six measurements per sample were averaged to obtain the sample's grain size distribution (Vogel et al., 2016). Particle sizes are defined according to Ad-hoc Arbeitsgruppe (2005) and reported in vol%.

**Element Concentrations** Analysis of element concentrations was conducted with a Niton XL3t portable energy-dispersive X-ray fluorescence spectrometer (p-ED-XRF) by Thermo Scientific (Waltham, Massachusetts, USA). Internal parameter calibration was set to 'mining mode' and all

measurements were recorded in %. Two certified reference materials (CRMs) were used for quality control: GBW07312 (stream sediment; National Research Center for Certified Reference Materials in China) and NCS DC 73387 (soil; China National Analysis Center for Iron and Steel). The CRMs were measured repeatedly after every eight samples. Since this study does not examine any absolute elemental concentrations but, rather, the overall variability throughout the profile, the CRMs were used only to ensure validity of the measured data. Recovery values were therefore not examined quantitatively.

Further processing of the p-ED-XRF data was necessary to produce reliable measurement results. Measurements with a value smaller than or equal to three times the $1$-$\sigma$ error were discarded. Following Schmidt et al. (2019), the remaining results were analysed using a compositional data approach to manage the consistent sum-constrained model (Aitchison, 1986) via R (R Core Team, 2018). Elements were only processed if the number of values below the limit of detection did not exceed 10% of the respective measurements. Subsequently, the data were transformed using centred log-ratio (clr; Aitchison, 1986):

$$z = clr(x) = [ln(x_1/g(x)), ln(x_2/g(x)), ..., ln(x_D/g(x))],\qquad(5.1)$$

where $g(x) = [x_1 \times x_2 \times ... \times x_D] \times 1/D$ is the geometric mean of all parts of the observation $x = [x_1, x_2, ..., x_D]$. This transformation was conducted using the compositions package for R (Boogaart et al., 2018). Prior to the conducted cluster analysis, the data were additionally standardised using a robust z-transformation.

**Visible and Near-Infrared Spectroscopy**   Visible and near-infrared reflectance of the sediment samples was captured using an FieldSpec II spectroradiometer by ASD (Malvern, Worcestershire, UK). Samples were measured before and after homogenisation. This allowed us to indicate an influence of material, which may be coating quartz grains and, therefore, over-represented in the unground samples (e.g., fine clay or Fe oxides, see Soriano-Disla et al., 2014), and likewise to assess the general influence of grinding and homogenisation on sediment colour (Sánchez-Marañón et al., 1995). Reflectance was measured between 350 and 2500 nm in 1 nm steps. Measurements were acquired in a darkroom, where each sample was uniformly illuminated by two halogen lamps to minimise the casting of shadows. Samples were spread homogeneously on a plate, and the fibre optic cable of the spectroradiometer was placed above the samples (Online Resource 5.3 in Appendix). Each spectrum was averaged from 50 single measurements to compensate for uneven sediment texture. A white reference was captured every 15 min to minimise variation between the measurements and to acquire normalised reflectance values in the range [0,1].

The raw spectral information showed a high noise content below 450 nm and above 2300 nm, as well as several smaller aberrations. The spectra were therefore smoothed using a Savitzky–Golay filter of second polynomial order with a width of 21 values (Savitzky and Golay, 1964). This was conducted with the hsdar package for R (Lehnert et al., 2017). Due to the high amount of noise below 450 nm, the spectral data were limited to the range between 450 and 2500 nm for further processing (Figure 5.4).
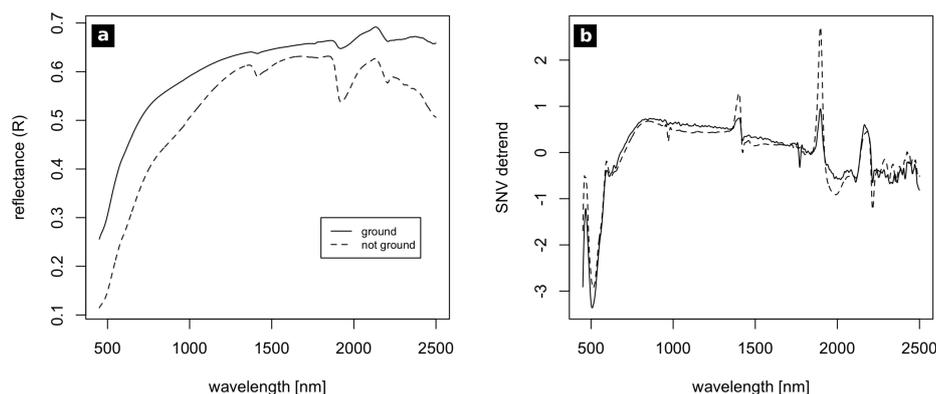
**Figure 5.4:** Exemplary spectral signal (450–2500 nm, sample no. 10) before (**a**) and after (**b**) pre-processing. Standard normal variate (SNV) transformation with detrending was applied to the first derivative of the $log(1/R)$ transformed spectrum.

Several studies recommend analysing spectral absorbance rather than spectral reflectance (e.g. Viscarra Rossel et al., 2006; Stenberg et al., 2010). Spectral reflectance R was therefore transformed into absorbance $A = log(1/R)$, which allows for a better correlation of the signal with sedimentological data.

As an increase of overall reflectance was observed after homogenisation of the samples, several processing steps were carried out to compensate for this (Figure 5.4). Building upon Stenberg et al. (2010), baseline correction was conducted by calculating the first derivative, which eliminated the existing baseline shift (Owen, 1995). Furthermore, this step enhanced weak spectral signals. Since this step also introduced additional noise, the baseline corrected spectra were then smoothed using a Savitzky–Golay filter of second polynomial order with a width of 21 values. Subsequently, scatter correction was performed by a standard normal variate (SNV) transformation and de-trending (Barnes et al., 1989). These pre-processing steps allow for the direct comparison of the spectra of homogenised and non-homogenised samples as well as an enhancement of relevant spectral peaks and a reduction of, e.g., overall curvature (Stenberg et al., 2010). A more detailed documentation of the workflow is delivered in Online Resource 5.3 (Appendix).

### 5.2.2.2 Image Acquisition

Digital RGB photographs of section SD17P1 were taken using a 24.3 MP mirrorless camera with a 2.8/16 mm lens (Sony ILCE-6000, Sony SEL-16F28, by Sony, Tokyo, Japan). Multispectral imaging was carried out using a Micro-MCA snapshot camera with six spectral channels (Table 5.1) by Tetra-Cam (Chatsworth, California, USA). Central wavelengths and FWHM (full width half maximum) of the Sony ILCE-6000 RGB camera have been previously published by Haburaj et al. (2019) and used throughout this study.

Using optical measurement systems, the lighting of the profile needs to be controlled and documented. Images were captured under natural diffuse daylight. Apart from smaller diffuse shadows in the corners, the section was captured under uniform lighting conditions. The cameras were placed on a tripod and centred in front of the respective section. Rotation of the cameras was limited to

**Table 5.1:** Technical specifications of the camera systems used; central wavelengths of the spectral channels blue, green, red, red edge, NIR1, and NIR2.

|  | Sony ILCE-6000 | TetraCAM MicroMCA-6 |
|---|---|---|
| B | 485 nm | 490 nm |
| G | 540 nm | 550 nm |
| R | 595 nm | 680 nm |
| red edge | - | 720 nm |
| NIR1 | - | 800 nm |
| NIR2 | - | 900 nm |
| spatial resolution | 5696 × 4272 px | 1280 × 1024 px |
| colour depth | 8 bit | 10 bit |

a minimum, and black-and-white reference targets were placed in the section, which allowed us to overlay and merge images manually afterwards while keeping spatial distortions at a minimum.

The RGB images were converted into 8-bit TIF format and normalised to the range [0,1] via feature-scaling. Six multispectral images were recorded in TetraCam RWS-format (each containing one spectral channel and stored in three 8-bit layers, labelled red, green, and blue) and, to allow further processing, converted into six 10-bit TIF-format images using the formula provided by the manufacturer.

Since the used multispectral camera consists of an array of six separate cameras and the images were acquired from a distance of approximately five metres, a significant spatial offset between the spectral channels was present and required manual band alignment. The proprietary processing software PixelWrench2 contains an option for minimising this offset, which was used successfully by, e.g., Retzlaff (2015). This solution, however, is more suitable for images acquired during remote sensing, as they keep a certain minimal distance from the subject, therefore keeping the spatial offset at acceptable levels. Our ground-based approach, on the contrary, included an offset that was too high and, therefore, was corrected through the pairwise manual selection of reference points. The resulting multispectral images were normalised to the range [0,1] via feature scaling.

### 5.2.2.3 Image Analysis

**Image Pre-Processing**    To allow for the comparison of the classification results via spatially consistent image data, the recorded RGB and multispectral images were processed by manually registering 12 reference points (overlapping geometries and artificial black-and-white targets) for thin plate spline (Duchon, 1977) transformation (TPS) with nearest neighbour resampling, carried out with the Georeferencer plugin in QGIS (v2.18). This resulted in overlying images, which are necessary to assess the quality of the final results. All non-sediment objects captured in the images were masked manually. Aside from external disturbances like grass patches and other vegetation, this included the prehistoric stone layers that are separating the sediment layers from one another (Figure 5.2). This step was conducted using the raster package for R (Hijmans, 2019). Minimising the impact of image noise and smaller disturbances within the profile, the images were filtered spatially using the raster package for R (Hijmans, 2019). To keep the geometric alteration as low as possible and,

likewise, to minimise the influence of outliers, we chose a median filter with a window size of 21 by 21 pixels (multispectral data) and 101 by 101 pixels (RGB data) for spatial filtering. The window size was scaled according to the respective spatial resolution and set to quite high values, since our main interest was in the main strata. The sediment layers examined did not include flakes of, e.g., brick, charcoal, or mortar, which renders the loss of texture detail in the image data acceptable (Haburaj et al., 2019).

**Slope Rasters**   Various derivatives of raster data show a positive impact on classification results. According to Haburaj et al. (2019), the spectral slope between red, green, and blue channels—while introducing more noise—can help to identify organic material in the images. Slope rasters were therefore created for both the RGB and the multispectral image data, leading to two and five additional raster layers, respectively.

**Principal Component Analysis**   Additional raster layers were created by performing a principal component analysis (PCA, conducted via Leutner and Horning, 2017) on the RGB image and the multispectral image. The number of principal components was set to two, since this delivered an acceptable explained cumulative variance (Online Resource 5.1 in Appendix).

**CIELAB**   The colour spaces CIELAB, CIE-Yxy, and Munsell are suited for the analysis of iron oxide content in soils and sediments (Jarmer and Schütt, 1998; Schwertmann and Cornell, 2008). The CIELAB colour space shows a clear benefit regarding quantitative analyses and was used throughout the presented study, since CIELAB coordinates are often influenced by soil and sediment properties: CIELAB lightness L* is often regarded to be related with organic matter, and CIELAB chromaticity coordinates a* and b* were proven to be related with different iron oxide concentrations (cf. Barron and Torrent, 1986; Jarmer and Schütt, 1998; Vodyanitskii and Kirillova, 2016)). While these relations are measurable, authors generally conclude that the specific relation is highly dependent on the material examined and often cannot be transferred between different soil types (Vodyanitskii and Kirillova, 2016).

The recorded image data were therefore transformed into the CIELAB colour space using the convertColor() function for R. The CIE standard illuminant D65 was used as reference white, since its properties are close to real indirect daylight as present during fieldwork. The multispectral data were modified so that only the red, green, and blue bands were transformed. The produced raster layers were used as additional input for classification with the intent of adding a rough measurement of lightness and, more importantly, iron oxide content.

**Soil Organic Carbon**   Numerous studies have examined the prediction of soil organic carbon (SOC) via the spectral signal (Baumgardner et al., 1985; Hill and Schütt, 2000; Mutuo et al., 2006; Brown et al., 2006; Sankey et al., 2008; Volkan Bilgili et al., 2010; Steffens et al., 2014; Hobley et al., 2018). Most recent publications tend to stress the better accuracy of machine learning algorithms (Stenberg and Viscarra Rossel, 2010), especially when predicting more complex soil properties. For the prediction of SOC, however, linear regression models have proven to deliver robust predictions

with acceptable accuracy. Since our study solely aims for the rough delineation of strata and our number of prediction variables is quite low, ordinary least squares (OLS) regression was chosen as a rather straightforward prediction model. As already pointed out by Baumgardner et al. (1985), the overall reflectance in the visible part of the spectrum decreases when the organic matter content increases. For this reason, only the first four image bands of the multispectral image were used as input data for the prediction of SOC (450, 550, 680, and 720 nm). SOC measurement results were assigned to a set of training pixels, which were extracted from the image bands based on the respective sampling locations. Training pixels comprise all pixels of the digitised sample location. As the number of pixels is quite small due to the limited spatial resolution of the images, further processing of the pixel values via, for example, mean values of random subsets (as suggested for bigger data sets by, e.g. Hobley et al., 2018) was not conducted. The fitted model was used for prediction of SOC based on all image pixels. Model fitting was performed with the stats package for R (R Core Team, 2018) and prediction using the raster package for R (Hijmans, 2019).

### 5.2.2.4 Cluster Analysis

Unsupervised classification of the processed data was conducted to semi-automatically delineate stratigraphical (sub)units from (i) the image data and their derivatives and (ii) multiple parameters of the sedimentological data. This type of classification does not require the definition of training areas and, thus, assumptions prior to the performance of the classification; it therefore excludes the researcher's subjective perception (Haburaj et al., 2019). The on-site delineation by local archaeological experts in combination with the results of the quantitative sedimentological analyses was used as a reference to discuss the quality of the results.

**Sedimentological Data**    Unsupervised classification of the sediment properties examined was used to verify stratigraphic delineation conducted during fieldwork. The most recent work of Schmidt et al. (2019) shows that hierarchical clustering of element concentrations may be used to classify sediment data for stratigraphic delineation and interpretation. Building upon their work, we applied cluster analysis of Manhattan distances performed with Ward's method (e.g. Strauss and Maltitz, 2017). Additionally we examined the benefit of depth-constrained clustering (Gill et al., 1993) using CONISS (Grimm, 1987). The number of clusters/groups is based on the layers identified in Figure 5.2c: Four main layers (A, B, C, D) and two additional layers, representing the sediment parts within the stone layers. The total number of six clusters was increased to 10 clusters for depth-constrained clustering, now including transitional layers between the main layers. Both clustering approaches were conducted for (i) the pre-processed element concentrations (p-ED-XRF), the pre-processed spectra of the (ii) ground and (iii) unground sediment samples, and (iv) the median pixel values of the RGB and (v) multispectral images, extracted from the sampling locations.

**Image Data**    In past studies, k-means clustering was successfully applied for the analysis of soil data (e.g. Steffens et al., 2014; Zhang and Hartemink, 2019b; Haburaj et al., 2019). In this study, clustering was carried out in the spectral domain using the k-means algorithm proposed by Hartigan and Wong (1979). Partitioning into clusters was performed using a random sample, consisting of 10%

of the pixels of the respective raster stack (R Core Team, 2018; Hijmans, 2019) in order to reduce computational time. The maximum number of iterations allowed was set to 50, and three random sets were created as seed points. The resulting partition was used for prediction via the clue package for R (Hornik, 2005).

Following Haburaj et al. (2019), multiple combinations of the processed image data were used as input for the k-means clustering: All image derivatives created were combined with one another to assess the influence of each processing step regarding image noise and cluster quality. All combinations examined are presented in the supplemental material (Online Resource 5.1 in Appendix).

The number of clusters was set to 15, allowing the inclusion of potential disturbances and transitional layers in comparison to the six clusters of the sedimentological data. This rather high number of clusters leads to the necessity of subsequent manual grouping of the clusters with respect to the on-site stratigraphic documentation. This step was conducted in QGIS (v2.18). As argued by Haburaj et al. (2019), manual grouping should be conducted attentively, as it allows for a more transparent documentation of the where and how of drawing borders between layers.

Evaluation of the image clustering results was carried out manually and by a visual interpretation of the homogeneity of the resulting clusters, as well as their conformity with the delineation of stratigraphic layers depicted in Figures 5.2 and 5.5.

## 5.2.3   Results

### 5.2.3.1 Sedimentological Record

The measured sediment properties of the section examined allow a detailed description of the layers depicted in Figure 5.2. Layer A is characterised by a water content of c. 5%, an average SOC concentration of around 1%, acidic pH values (top A: < 3.5, basis A: 4.5), and high concentrations of silt compared to the other layers (Figure 5.5). At the top of layer B (sublayer 5), a peak of water content is visible (c. 10%). Layers B and D show strong variation between their individual sublayers, averaging at around 5% water content, acidic pH values between 3.5 and 4.5, and low SOC concentrations (<0.2%). The dark, thin horizontal layer C shows a peak in water content (c. 10%) and a slight increase of silty material and SOC content compared to the surrounding materials (Figure 5.5).
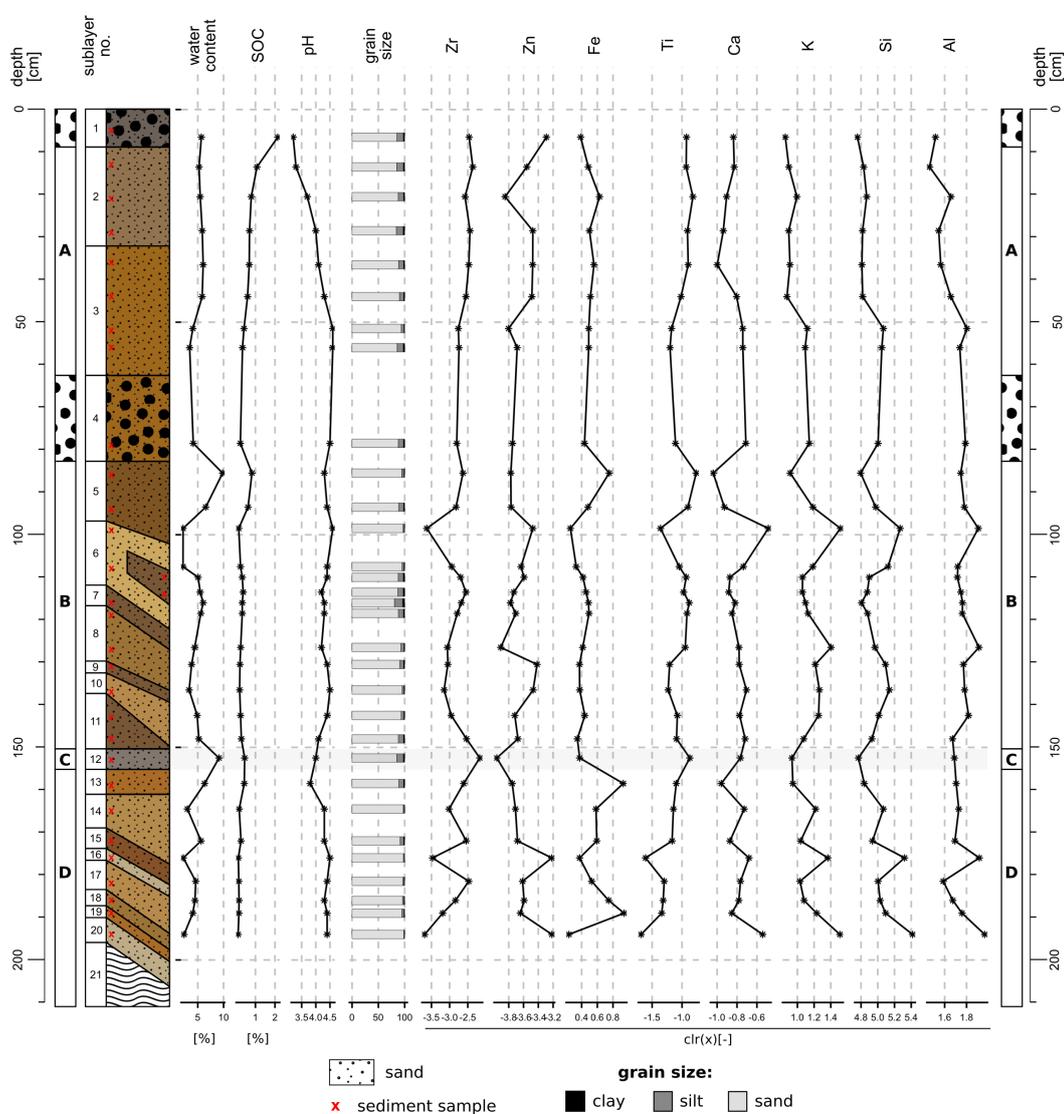


**Figure 5.5:** Sedimentological record of SD17P1 (layers A, B, C and D). Sediment colour and texture (Schoeneberger et al., 2012) as documented during fieldwork, location of samples, and respective results of the sediment analyses performed. Centred log-ratio (clr) transformed contents of relevant elements measured with the p-ED-XRF device.

During pre-processing of the p-ED-XRF data, we excluded several elements due to their insignificance or erroneous measurements. The following eight elements showed valid measurements and were processed: Zr, Zn, Fe, Ti, Ca, K, Si, and Al. Concentrations of these elements are heterogeneous in the lower parts of the examined profile (layers B, C, D), while layer A repeatedly shows continuous trends of element concentrations from its top to its bottom (Figure 5.5).
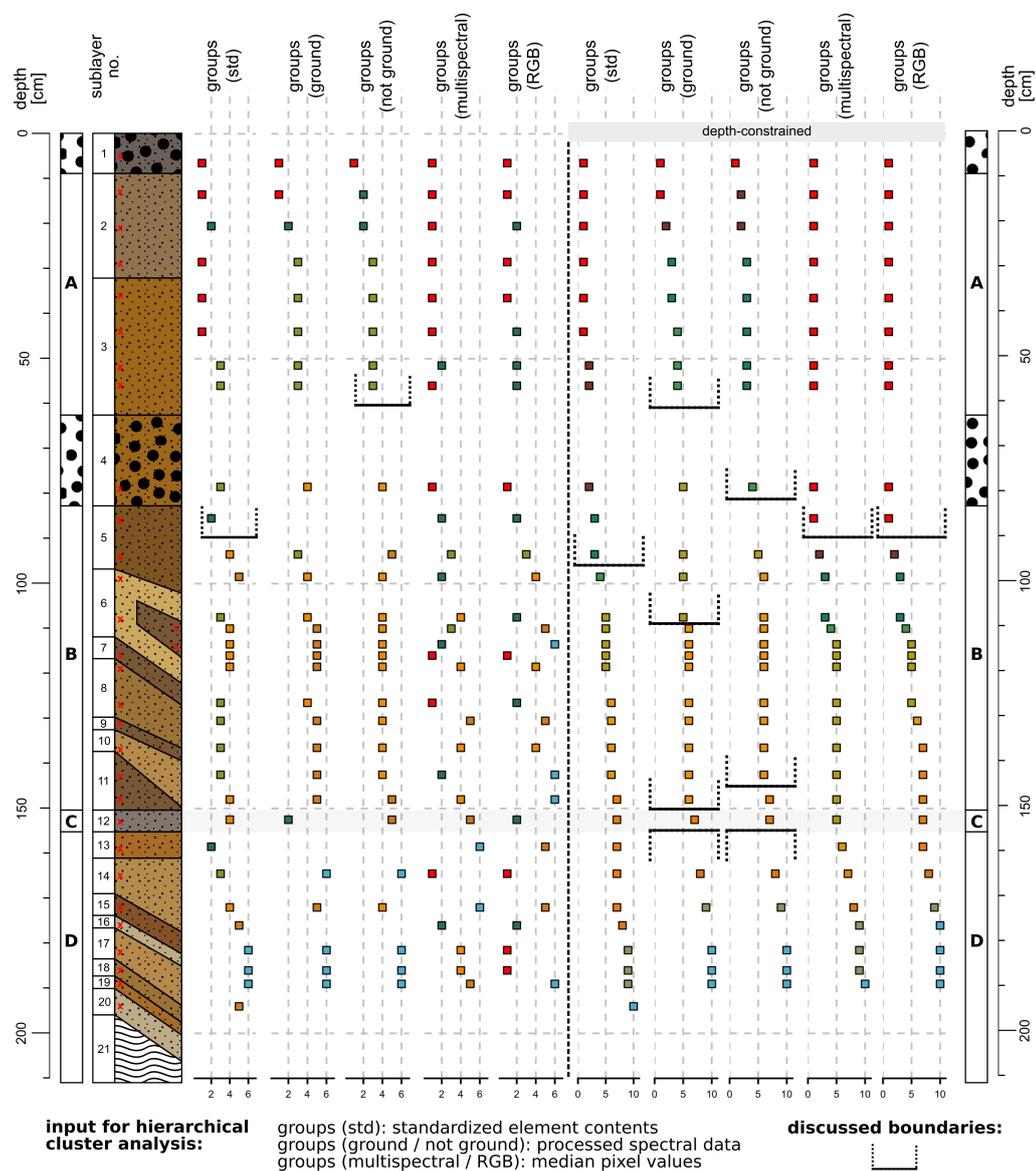


**Figure 5.6:** Clustering results (groups) of SD17P1 (layers A, B, C and D). Sediment colour and texture (Schoeneberger et al., 2012) as documented during fieldwork, location of samples, and respective results of the hierarchical and the depth-constrained hierarchical cluster analyses of the following data: p-ED-XRF data, selectively measured spectral data (450–2500 nm), and median pixel values extracted from image data. Colours of samples indicate groups obtained from the cluster analyses. Six groups (six colours) were obtained from the hierarchical clustering and ten groups (ten colours) from the depth-constrained hierarchical clustering. Boundaries discussed in the text are highlighted.

### 5.2.3.2 Clustering Results – Sedimentological Data

Clustering results (groups) of the p-ED-XRF data are documented in Figure 5.6. The groups produced by depth-constrained hierarchical cluster analysis of the transformed element concentrations capture the main stratigraphic layers of the profile and, together with the on-site delineation and the sedimentological record (Figure 5.5), serve as a reference for the quality of all clustering results: Groups 1 and 2 cover areas of dark material, which are characterised by high water contents and are situated in layer A and at the top of layer B (sublayer 5, see Figures 5.5 and 5.6). Layer C is grouped together with the surrounding materials of layers B and D in the results of both clustering approaches.

Clustering of the pre-processed spectra of the samples showed a noteworthy influence of grinding. While layer C is grouped with the topmost layer in the ground data, it is grouped with the darker part of layer B when samples are not ground (Figure 5.6). Both ground and not ground samples clearly delineate data of layer A from the rest of the data in the clustering results. Layer C is assigned to a separate group in the ground data only. Compared to the results of the p-ED-XRF data, the groups produced on the basis of the spectral data are spatially more consistent. Depth-constrained clustering increases this consistency (Figure 5.6).

Groups produced by the clustering of both RGB and multispectral pixel data capture layer A homogeneously. Additionally, the hierarchical clustering assigns darker parts of layers B and D to groups 1 and 2. The clustering of the RGB data groups layers A and C, while the clustering of the multispectral data groups layer C with the darker parts of layer B. Both data sets assign the topmost part of layer B (sublayer 5, see Figure 5.6) to groups 1 and 2. Including depth-constrained clustering showed a positive effect on group homogeneity for both RGB and multispectral-based clustering results; layer A and sublayer 5 are captured as one single group. Both depth-constrained cluster approaches of the pixel data, however, do not capture layer C, but group it together with the surrounding materials of layers B and D.

### 5.2.3.3 Clustering Results – Image Data

Multiple band combinations were tested as input for the performed image classification via k-means clustering. While all combinations have problems capturing the thin horizontal layer C, the results shown in Figure 5.7 indicate clear benefits when using specific bands: Using the RGB data, the most homogeneous classification results were derived from a combination of the raw RGB bands, two principal components, and two slope rasters (Figure 5.7b). The summarised groups produced by the cluster analysis roughly represent the stratigraphic (sub)layers. Clustering the same band combination of the multispectral data (Figure 5.7c) produced results of similar quality. Nevertheless, two major differences have to be pointed out: (1) Multispectral data succeed in separating layer A from the rest of the profile, and (2) RGB data capture layer C more precisely in the spatial domain. Layer C is grouped with materials from layer A in the results produced using RGB data, while the multispectral data clustering results group it together with parts of layers B and D (see Online Resource 5.1 in Appendix for all band combinations).
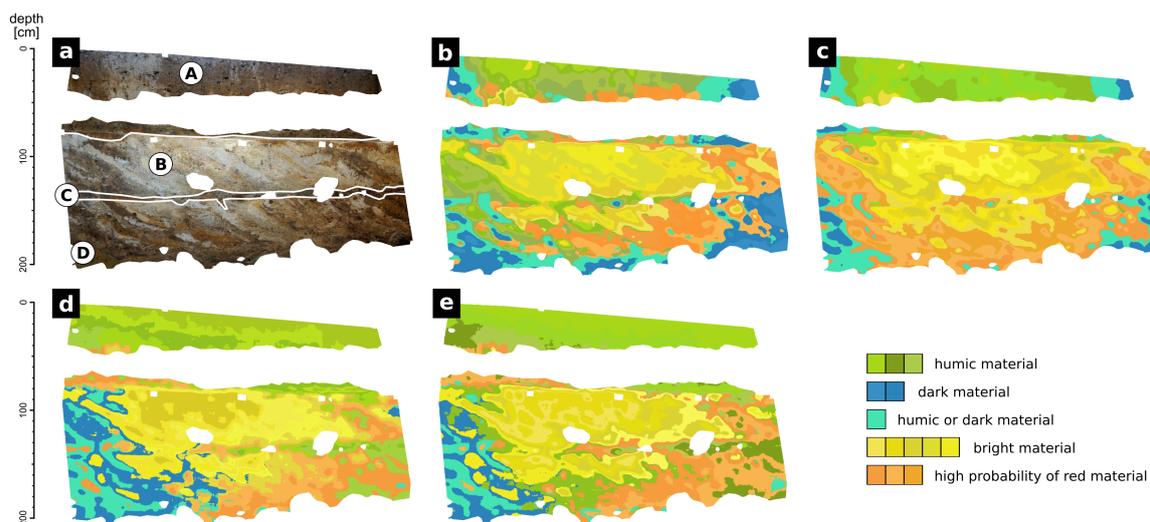
**Figure 5.7:** Results of the performed image classification of the upper part of SD17P1; stone layers were masked manually prior to processing. (**a**) Profile as perceived during fieldwork with layer borders (white) according to Figure 5.2 (layers A, B, C and D). Classification results of (**b**) an RGB composite (raw data, slopes, PCA results), (**c**) a multispectral composite (raw data, slopes, PCA results), (**d**) a multispectral composite (slopes, CIELAB), and (**e**) a multispectral composite (slopes, predicted SOC, CIELAB chromaticity coordinates a* and b*). Colour shades represent individual groups produced from k-means clustering, while summarised groups are indicated by a common hue.

## 5.2.4 Discussion

### 5.2.4.1 Image Data

Our image classification results of the RGB data are consistent with the results of other studies (e.g. Zhang and Hartemink, 2019b,c,a; Haburaj et al., 2019). The combination of raw RGB bands, two principal components, and two slope rasters (as proposed by Haburaj et al., 2019) produced similar results to those of the equally processed multispectral bands. Furthermore, processing of the multispectral data improved the classification results significantly, allowing us to minimise the impact of lighting differences successfully.

Transformation to CIELAB coordinates did not deliver any benefit for the RGB data: Neither replacing the RGB bands with their CIELAB counterparts, nor using the image bands containing chromaticity coordinates a* and b* separately led to better clustering results (Online Resource 5.1 in Appendix). This is different for the multispectral data, as both the CIELAB data based on the multispectral data and SOC data show a significant influence on the clustering results (Figure 5.7d,e). A combination of slope bands and CIELAB bands eliminates the influence of slightly shadowed areas on the left and right sides of the profile. This leads to more homogeneous results, which is mostly visible in layer A and the righthand parts of layers B and D. As a consequence layer C is—while grouped with the material of layer A—captured more reliably than with the other multispectral band combinations. Another notable result is the similar performance of (i) the combination of slope bands with all CIELAB bands (Figure 5.7d) and (ii) the combination of slope bands with the CIELAB chromaticity bands a* and b* and the SOC raster (Figure 5.7e). Classification results are nearly identical, only showing small differences in the lower part of the profile.

We observed similarities between the clustering results of band combinations including the SOC raster and, alternatively, the CIELAB lightness raster. Nevertheless, the transferability of these results to other sites has to be examined critically and individually for each case. The noticed relation between SOC and lightness rasters is already proposed by other authors (e.g. Barron and Torrent, 1986; Vodyanitskii and Kirillova, 2016), who examined this phenomenon in detail and concluded that the relation is highly dependent on the local circumstances. However, if quantitative examination of SOC as conducted by, e.g., Steffens et al. (2014) cannot be included into the workflow due to technical or financial limitations, we propose CIE lightness to be a rough indicator of the spatial variance of SOC, while keeping in mind that water content also influences lightness significantly (e.g. Baumgardner et al., 1985; Stenberg et al., 2010). Our results support the strong correlation of CIELAB image bands and SOC, as observed for high-contrast soil profiles, for example, by Zhang and Hartemink (2019a).

The initially proposed significance of CIELAB chromaticity coordinates for the examination of iron oxide concentrations could not be confirmed for our experimental design. The examination of iron oxides based on CIE a* and b* as proven by, e.g., Vodyanitskii and Kirillova (2016) possibly requires a more sophisticated approach, including the mathematical prediction via, e.g., partial least squares regression or machine learning algorithms, as applied for other sediment properties by, e.g., Daniel et al. (2003), Viscarra Rossel and Behrens (2010), or Morellos et al. (2016). Other factors that could be responsible for the poor performance of our approach include the limited presence of iron in the profile examined: Elemental Fe concentration is below 1% (see Online Resource 5.2 in Appendix) throughout the examined samples. This would correspond to a maximum of c. 1.4% $Fe_2O_3$, given that all Fe would result from hematite. A future approach should consider a camera setup with fewer limitations regarding spectral range and resolution. Using a hyperspectral VIS-NIR device would allow the examination of Fe in far more detail (Baumgardner et al., 1985; Viscarra Rossel et al., 2009; Stenberg and Viscarra Rossel, 2010). This would also allow the analysis of further sediment properties, like grain size based on the spectral signal (Post and Noble, 1993; Bishop et al., 1994; Madejova and Komadel, 2001; Stenberg et al., 2010).

More generally speaking, our results support the usage of the CIELAB colour space as a standard for documentation during archaeological excavations. The observed good performance of the CIELAB datasets supports the scientific relevance of the CIELAB colour space, which was previously proposed by other authors (Sánchez-Marañón et al., 1995; Fairchild, 2013). For the last decades, this colour space gained popularity in the geosciences and, meanwhile, is widely used in numerous studies (cf. Debret et al., 2011). Many authors noted the clear advantages over the RGB or Munsell colour systems (e.g. Melville and Atkinson, 1985; Viscarra Rossel et al., 2006; Sprafke, 2016): CIELAB is a mathematically defined colour space that is free to use and suitable for quantitative analyses. In future studies, the proposed quantitative examination of colour differences throughout archaeological profiles by transformation to CIELAB could be complemented by additional colour difference formulas like CIEDE2000 to assess possible benefits (Liu et al., 2013).

### 5.2.4.2 Sedimentological Data

By quantitatively analysing sediment properties, we were able to explain several characteristics of the profile in detail. Clustering results suggest that the topmost parts of layer B (sublayer 5) are related to layer A or the intermediate stone layer. The observed high values of water content and SOC suggest sublayer 5 to be either related to construction work (compression by overlying stones) or resembling (together with the overlying stones) a prehistoric surface, which would have been of similar character to the topmost stone layer described by May (2018a). Similarly to sublayer 5, the colour of layer C seems to be influenced mainly by a high water content. However, slightly increased SOC may also indicate a prehistoric surface here. The observed similarities also explain why layer C is often clustered with layer A or sublayer 5.

The performed hierarchical cluster analysis of the examined sediment properties produced similar results for the p-ED-XRF data and the recorded VIS-NIR data (Figures 5.5 and 5.6). We see this as an indicator that selective or extensive spectroscopy should be considered a fast and efficient alternative to laboratory analyses when one is aiming for the delineation of layers or an objective confirmation of the observed variance in a profile. Spectroscopy in comparison to laboratory analyses is easy and fast to apply, non-destructive, cost-efficient in the long term, and also more environmentally friendly. While our study supports the idea of using spectroscopy during archaeological excavations due to the aforementioned benefits, the methodological approach should still be considered experimental—currently, clear limitations regarding the ease of use and the on-site applicability exist.

Throughout our study, depth-constrained clustering generally improved the results, which supports preferability of this method when clustering layers, as, e.g., Schmidt et al. (2019) proposed. Depth-constrained clustering also improved clustering results of the sampled pixel values significantly (Figure 5.6) and—with the exception of layer C—captured overall variability of the profile. Layer C is generally treated similarly in the image data and the selective data, as it is either clustered with layer A or darker parts of layers B and D. It is also noteworthy that layer C was captured most accurately by the spectral data of ground sediment samples, which underlines the potential of VIS-NIR spectroscopy for stratigraphic analyses. In addition, both the selective pixel-based approach and the extensive image-based approach distinguish between layer A and the rest of the profile.

### 5.2.4.3 Shortcomings and Benefits of Spectral Measurements

Our study exposes that quantitative colour measurements during fieldwork can help with stratigraphic delineation and interpretation, be it based on selective or extensive data. Combining these measurements with semi-automated classification algorithms, we were able to increase objectivity and transparency of the delineation of stratigraphic (sub)layers in a similar manner to that if using common sedimentological data. In particular, depth-constrained hierarchical clustering should be considered a useful tool when analysing stratigraphic sequences quantitatively. Depending on the respective research question, future studies should consider the measurement of VIS-NIR spectral data as an alternative to other analytical methods, which are often time-consuming and destructive. For example, spectral measurements could offer a fast and quantitative way of documentation on rescue excavations.

We were able to successfully transfer the image analysis workflow proposed by Haburaj et al. (2019) to the archaeological profile SD17P1 from the burial mound of Seddin. Moreover, we want to stress the significant improvements we could add to this process by adding CIELAB data and SOC information, both derived from the multispectral data. The observed similarities between the clustering results of band combinations, including the SOC raster and, alternatively, the CIE lightness raster, suggest that time-consuming and expensive SOC prediction can be replaced with a fast transformation of the image data to the CIELAB colour space. The noticed benefits should especially be kept in mind when one is working under heterogeneous fieldwork conditions, as CIELAB and slope rasters clearly help to overcome uneven lighting conditions. If intra- or inter-site comparability of the colour and spectral measurements is required, image acquisition should include distinct colour calibration (Sánchez-Marañón et al., 2011; Funt and Bastani, 2014).

The output of the conducted comparisons suggests that the selective spectral information measured from sediment samples produces similar clustering results to those of the p-ED-XRF data. Increasing the spectral range and resolution of the image data would therefore clearly improve the identification of (sub)layers based on image data. This is consistent with the results of laboratory analyses of, e.g., Steffens et al. (2014) or Hobley et al. (2018).

Taking the characteristics of the examined profile into account, we consider its stratigraphy complex: While it consists only of a few major layers, the overall heterogeneity of the profile is quite high. This complexity renders sediment sampling rather complicated, as one has to decide which and how many samples to take. At this point, analyses based on image data could fill a gap. This would, however, require a device that offers an improved spectral range and quality. Spectral imaging devices are still in a crucial phase of development though. Imaging technology becomes more sophisticated: Systems become smaller and are more easy to use, as recent studies suggest, using either snapshot (e.g. Bareth et al., 2015; Haburaj et al., 2019) or push broom (e.g. Behmann et al., 2018) technology. However, the offered devices are still often limited to wavelengths below 1000 nm (cf. Lodhi et al., 2018), rendering the examination of many sediment properties problematic. Established hyperspectral scanners that capture wavelengths beyond 1000 nm are most often expensive and difficult to use under fieldwork conditions due to their dimensions, weight, and power consumption. While the limited capabilities of the presented spectral imaging devices render the extensive analysis of sediment data not yet applicable for most excavations, selective spectral measurements as performed in this study may constitute an easy-to-apply additional step towards a more transparent way of field documentation during archaeological excavations.

Keeping in mind the complex stratigraphy of the profile at hand, applicability of the proposed approach to profiles and plana that include more distinct stratigraphic features (e.g., post-holes, pits, trenches) appears likely. The good performance of RGB image analysis for the examination of high-contrast soil and sediment data was lately described by Haburaj et al. (2019) and Zhang and Hartemink (2019a), indicating a high potential of image classification for archaeological excavations in general. Our results of the selectively measured colour and spectral data indicate that hyperspectral measurements (as we measured with a spectroradiometer) could offer supportive information when analysing more complex or low-contrast stratigraphies (e.g., paleosols).

The sedimentological record obtained in the laboratory is still necessary to describe and interpret

the layers. However, the good clustering results of the image data and the general performance of the selective spectral data stresses the potential of more advanced spectral image sensors for future stratigraphic studies. By extrapolating sediment properties from samples to image data, one could combine the benefits of both worlds: A sedimentological record that is not limited to selected sediment samples. While several laboratory studies were already able to produce extensive parameter maps (e.g. Steffens et al., 2014), transferability of this approach to on-site recordings still has to be examined.

Based on our results, future studies should either focus on the optimisation of an RGB-camera-based workflow or the in-depth review of more potent spectral sensors. While RGB images delineate layers in the presented study, our results show that image data that would involve spectral information of the same quality as our selective VIS-NIR measurements could prove to be as suitable for assisting stratigraphic delineation and interpretation as labour-intensive laboratory analyses.

## 5.2.5   Conclusions

The presented results indicate that quantitative colour measurements and spectral measurements during fieldwork can support stratigraphic delineation and interpretation of profiles in archaeological excavations. Depth-constrained hierarchical clustering of spectral data as well as of p-ED-XRF data should be considered a useful tool when analysing stratigraphic sequences quantitatively.

The discussed benefits of the multispectral image data suggest a huge potential of the spectral domain for automating stratigraphic delineation and interpretation and, likewise, render this process more traceable. Observed drawbacks show that we cannot offer a ready-to-use approach yet. In particular, the limited capabilities of the presented spectral imaging devices render the extensive analysis of sediment data not yet applicable for most excavations. Rather, additional on-site studies should be carried out to create standardised workflows. Image data that would involve spectral information of the same quality as our selective VIS-NIR measurements could prove to be as suitable for assisting stratigraphic interpretation as labour-intensive laboratory analyses. Quantitative colour measurements and spectral measurements, as performed in this study, may constitute an easy-to-apply, non-destructive additional step towards a fast and more traceable way of field documentation during archaeological excavations.

## Acknowledgements

CHAPTER 6

## On-site spectral imaging for predicting sediment properties

### 6.1 Introduction

The chemical and physical composition of soils and sediments heavily influences their spectral reflectance. Several studies explored this connection, which is often utilised to map parameters like organic matter or iron oxide contents based on spectral image data. Recent laboratory studies indicate a significant potential of this technique for the extensive digital mapping of soil and sediment profiles based on quantitative image data (e.g. Steffens and Buddenbaum, 2012; Zhang and Hartemink, 2019a).

With past studies being carried out mostly in the laboratory, the following issues need to be addressed: (i) can the mapping of sediment properties via spectral imaging be transferred from the laboratory to regular fieldwork conditions? (ii) How do the statistical models used perform when compared to stratigraphic documentation based on expert-knowledge? (iii) Can we also map complex archaeological profiles which include anthropogenic layers? And (iv) does the proposed workflow lead to reproducible results that render the act of stratigraphic analysis more traceable?

These topics are examined in the following study. For this purpose two archaeological profiles were recorded with a RGB and a hyperspectral camera. The acquired image data was used to train statistical models which involve physical and chemical sediment properties obtained from sediment samples analysed in the laboratory. The statistical models were used to produce reliable maps of soil organic matter, hematite, calcite, multiple weathering indices and grain size composition along the examined profiles. Validation of the produced results showed promising results for both imaging sensors, confirming the assumed high potential of optical data for stratigraphic documentation and interpretation.

## 6.2   Paper 3: Coupling spectral imaging and laboratory analyses to digitally map sediment parameters and stratigraphic layers in Yeha, Ethiopia

## Abstract

Quantitative analyses of soil and sediment samples are often used to complement stratigraphic interpretations in archaeological and geoscientific research. The outcome of such analyses often is confined to small parts of the examined profiles as only a limited number of samples can be extracted and processed. Recent laboratory studies show that such selectively measured soil and sediment characteristics can be spatially extrapolated using spectral image data, resulting in reliable maps of a variety of parameters. However, on-site usage of this method has not been examined. We therefore explore, whether image data (RGB data and visible and near infrared hyperspectral data), acquired under regular fieldwork conditions during an archaeological excavation, in combination with a sampling strategy that is close to common practice, can be used to produce maps of soil organic matter, hematite, calcite, several weathering indices and grain size characteristics throughout complex archaeological profiles. We examine two profiles from an archaeological trench in Yeha (Tigray, Ethiopia). Our findings show a promising performance of RGB data and its derivative CIELAB as well as hyperspectral data for the prediction of parameters via random forest regression. By including two individual profiles we are able to assess the accuracy and reproducibility of our results, and illustrate the advantages and drawbacks of a higher spectral resolution and the necessary additional effort during fieldwork. The produced maps of the parameters examined allow us to critically reflect on the stratigraphic interpretation and offer a more objective basis for layer delineation in general. Our study therefore promotes more transparent and reproducible documentation for often destructive archaeological fieldwork.

**Keywords:** archaeology; traceability; reproducibility; digital soil mapping; proximal sensing; random forest; stratigraphy; fieldwork

### 6.2.1   Introduction

In the past decade a growing number of institutions and authors postulate more transparent and reproducible research approaches (e.g. Mesirov, 2010; Peng, 2011). In this context, archaeologists began to critically reflect on multiple aspects of data generation, analysis and interpretation (e.g. Marwick, 2017; Marwick et al., 2017). Keeping in mind the often destructive character of archaeological excavations, transparency and reproducibility are highly relevant factors for archaeological fieldwork which involves on-site documentation and interpretation. One major part of this fieldwork is the delineation and characterisation of stratigraphic layers. While often accompanied by selective quantitative analyses of the samples extracted, the description and interpretation of the stratigraphic material remain in part influenced partly by the subjective perception of the respective researcher at work (cf. Haburaj et al., 2019). In the sense of Marwick et al. (2017), stratigraphic interpretation should therefore be based on a more open method, reproducible independent of the researcher's

experience. In the study at hand we present an approach that – in the long run – aims to promote a transparent, reproducible method of stratigraphic delineation and interpretation that is based on spectral recordings, i.e. physical information.

The often time-consuming quantitative analyses that traditionally support stratigraphic interpretation remain restricted to selective data based on the sampled materials and their laboratory analysis. By combining this data with digital image data, recent studies have successfully transferred selected soil and sediment properties, as derived from samples, to entire areas of soil and sediment cores and profiles: Steffens and Buddenbaum (2012) and Hobley et al. (2018), for example, were able to successfully map soil organic carbon (SOC) throughout soil and sediment profiles by analysing laboratory recorded hyperspectral data. These studies extend the subject of digital soil mapping (DSM) which until lately has been mainly concerned with the mapping of soil characteristics based on geodata like aerial or satellite image data (e.g. Hill and Schütt, 2000; McBratney et al., 2003; Lagacherie et al., 2006; Minasny and McBratney, 2016). Accordingly, digital maps of chemical and physical parameters throughout soil and sediment profiles could act as an additional layer of information for archaeological excavations and the interpretation of the excavated material. However, the studies of Steffens and Buddenbaum (2012) and Hobley et al. (2018) were carried out under controlled laboratory conditions. Transferability of the proposed method into a less controlled environment has been examined by Zhang and Hartemink (2019a) and Haburaj et al. (2020a), who show that RGB imaging and multispectral imaging are also suitable tools to map stratigraphic layers under fieldwork conditions. Zhang and Hartemink (2019a) examined the extrapolation of SOC content, pH values, grain size composition and weathering indices based on the RGB image data of a soil profile that was captured on-site. The resulting parameter-maps were highly accurate, which may be the result of their experimental design: a large number of samples (n = 90) were analysed for calibration from a small section (1.0 x 0.9 m) of a high-contrast Alfisol. Building upon these results, our study aims to predict the sediment properties of complex archaeological profiles from RGB and hyperspectral image data that were captured under fieldwork conditions. Our experimental setup features two archaeological sections (Figures 6.1, 6.2). The sampling strategy used for these sections follows common practice during excavations. The quality of the parameter-maps thus produced was assessed by comparison with stratigraphic delineation by archaeological experts (as depicted in Figure 6.2) as well as multiple statistical parameters. Our results show a promising performance of the proposed method for an extended stratigraphic analysis of archaeological excavations.

### 6.2.1.1 Study area and archaeological background

The settlement of Yeha is located 35 km to the northeast of Aksum and spreads over the eastern footslopes of several volcanic rock mountains (Japp et al., 2011; Gerlach, 2012). It is surrounded by a number of smaller periodically discharging channels which unite east of the settlement and subsequently traverse a fertile plain (Figure 6.1a). Besides a huge ancient settlement, two prehistoric monumental buildings are currently known within the area of the modern rural settlement of Yeha: the Great Temple in the east and the Grat Be'al Gebri in the north (Figure 6.1b). Both buildings are manifestations of a cultural transfer between indigenous societies located in todays northern Ethiopian highlands and the Sabaean society in South Arabia (Fattovich, 2012; Gerlach, 2013, 2015,
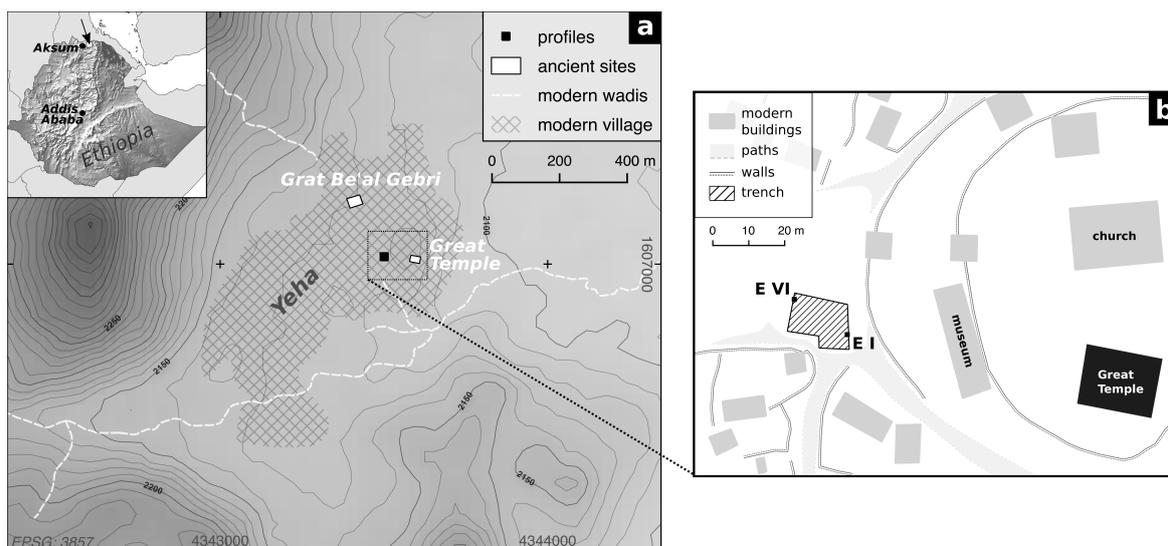
**Figure 6.1:** Topographical map of the study area (a) and the archaeological profiles examined (b). Elevation data: SRTM data (1 arc-sec.; U.S. Geological Survey).

2017, 2018) in the first half of the first millennium BC. While the character of these interactions is still under debate (Bent, 1893; Anfray, 1967; Fattovich, 1977; Phillipson, 2009; Nebes, 2010; Japp et al., 2011; Fattovich, 2012; Gerlach, 2013, 2015, 2017, 2018), there are obvious similarities in both the architectural and archaeological remains such as the use of the Sabaean language and South Arabian script, monumental architecture and building decoration, Sabaean deities and cultic rites, and bronze and other elements of the material culture. The results of archaeological research in various parts of the modern settlement suggest that the site was populated at least from the late 2$^{nd}$ millennium BC onwards (Japp, 2019), likely by sedentary farmers Brandt (1984). From the early 1$^{st}$ millennium BC onwards Yeha developed as the center of the Ethio-Sabaean community until, according to present knowledge, the mid-1$^{st}$ millennium BCE (Gerlach, 2017, 2018; Japp, 2019).

The profiles examined in this study were recorded during excavations by the Ethiopian-German Archaeological Mission to Hawelti, Yeha and surroundings (conducted as a cooperation project between the Sanaa Branch of the Orient Department of the German Archaeological Institute [DAI], the Seminar for Oriental Studies of the Friedrich Schiller-University of Jena, the Authority for Research and Conservation of Cultural Heritage [ARCCH] and the Tigray Tourism and Culture Bureau [TCTB]) at Yeha in 2018 and 2019. They are located in the area of the modern settlement of Yeha (Tigray, Ethiopia).

### 6.2.1.2 Profile description

Both profiles are part of an archaeological trench (Japp, 2019), found in the courtyard in front of the church compound; the Great Temple of Yeha is situated in the latter (Figure 6.1b). The examined areas of the profiles are shown in Figure 6.2. Profile EVI (western profile) covers several anthropogenic layers approximately 8 metres below the modern surface. The examined part of profile EI (eastern profile) involves several anthropogenic layers and a fine grained light brown horizontal layer which is present in many parts of the excavation but whose character is still under discussion. A detailed description of the two profiles EI and EVI is given in Table 6.1.
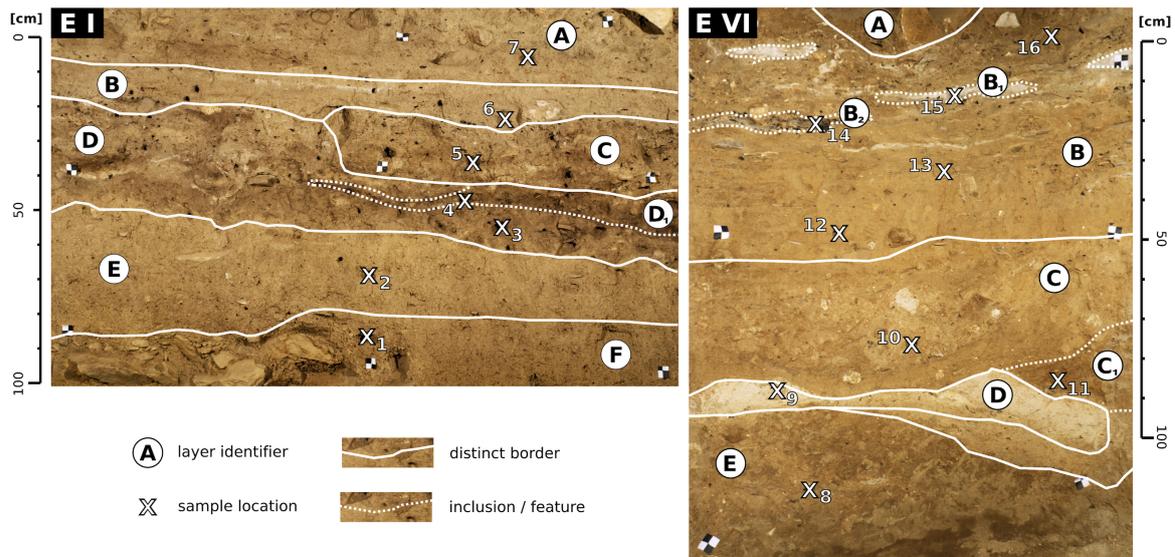
**Figure 6.2:** Archaeological profiles EI and EVI. The on-site stratigraphic interpretation is shown as white lines. Sediment samples were extracted at the 16 marked locations.

The two profiles do not overlap spatially and come from different depths of the archaeological trench. Additionally they date to different times: EI covers Aksumite and post-Aksumite times (1st half 1st to 1st half 2nd millennium CE), while EVI dates to the Ethio-Sabaean period with the lowest layer dating to the late 2nd millennium BC (cf. Japp, 2019). These characteristics lead to notable differences between the two profiles. There are perceptible variations of $PO_4$, $Fe_2O_3$, soil organic matter (SOM) and $CaCO_3$ contents throughout Profile EI (Figure A.1); otherwise EI is a low contrast profile with gradual change occurring in most of the examined parameters. The calculated weathering indices applied indicate minor differences between the layers examined in the present study but capture the overall structure of the stratigraphy (Table 6.1). Grain size composition only shows minor variation throughout profile EI. Contrary to this, Profile EVI shows more distinct differences between its layers (Figure 6.2, Table 6.1). Due to the strong variation of $PO_4$, $Fe_2O_3$, SOM and $CaCO_3$ contents as well as weathering indices and grain size composition, Profile EVI must be regarded as a high contrast profile with multiple layers that differ significantly in colour and overall brightness (i.e. spectral properties), the most prominent layers are EVI-B1, EVI-B2 and EVI-D. Detailed results of the sedimentological analyses carried out for profiles EI and EVI are shown in Table 6.1 and Figure A.1. Both profiles additionally differ in their situation within the trench: profile EI is part of a free-standing profile wall that is easily accessible whereas Profile EVI is situated in a narrow part of the trench where orthogonal profile walls are present directly left and right of the areas examined (Figure 6.2).

## 6.2.2  Materials and methods

Our methodological approach includes (i) the laboratory analysis of chemical and mineralogical components as well as grain size composition of the sediment samples, (ii) the acquisition of multi- and hyperspectral image data, (iii) random forest regression analyses and predictions of the examined sediment parameters with the spectral data received from the images, and (iv) accuracy assessment

**Table 6.1:** Stratigraphic layers of profiles EI and EVI. Selected sediment properties obtained from the sampled materials (KT: K-Ti-ratio; RX: Ruxton-ratio). Depth of samples given in cm below the modern surface. See Figure A.1 and Online Resources 6.1 and 6.2 for the detailed sedimentological record.

**Profile EI**

| Layer | sample no. | sample depth [cm] | Description | SOM [mass%] | $Fe_2O_3$ [mass%] | $CaCO_3$ [mass%] | KT | RX |
|---|---|---|---|---|---|---|---|---|
| EI-A | 7 | 184 | anthropogenic layer; grey-brown silty loam; inclusions of ash, charcoal, bones and ceramics | 4.7 | 9.3 | 3.9 | 37.1 | 6 |
| EI-B | 6 | 200 | anthropogenic layer; grey-brown silty loam; few inclusions; homogeneous | 5.1 | 10.2 | 4.3 | 30.4 | 5.4 |
| EI-C | 5 | 210 | anthropogenic layer; compacted brown-grey silty loam; inclusions of charcoal, stones, bones and ceramics | 2.2 | 8.9 | 6.2 | 31.7 | 5.9 |
| EI-D | 4 | 221 | anthropogenic layer; dark grey and brown silty loam with a high content of charcoal and ash; some bones and ceramics | 3.9 | 10.2 | 3.6 | 27.9 | 6.2 |
| EI-D1 | 3 | 228 | a concentration of darker and more brown silty loam inside layer D | 3.2 | 11.3 | 2.6 | 27.9 | 6 |
| EI-E | 2 | 240 | compacted silty loam yellow-brown sediment with very few inclusions; high loam content | 2.6 | 12.9 | 1.6 | 27.6 | 5.7 |
| EI-F | 1 | 257 | anthropogenic layer; compacted brown-grey silty loam with local clusters of small stones; high loam content; some bones and ceramics, some charcoal | 2.1 | 13.4 | 2.1 | 26.3 | 5.2 |

**Profile EVI**

| Layer | sample no. | sample depth [cm] | Description | SOM [mass%] | $Fe_2O_3$ [mass%] | $CaCO_3$ [mass%] | KT | RX |
|---|---|---|---|---|---|---|---|---|
| EVI-A | - | - | a thick layer of broken stones (supposedly deposited intentionally); separates the examined section from overlying stratigraphy | - | - | - | - | - |
| EVI-B | 13 | 682 | anthropogenic layer; dark brown-grey silty loam; inclusions of charcoal, ash, bones, ceramics, and burnt clay | 1.4 | 14.1 | 1.5 | 13.6 | 4.2 |
| EVI-B1 | 15 | 657 | grey to white bands of ash scattered throughout layer B; partly with charcoal | 3.5 | 8 | 28 | 12.9 | 5.6 |
| EVI-B2 | 14 | 670 | black charcoal layer; partly with burnt clay | 7.5 | 12.3 | 10.3 | 13.7 | 4.7 |
| EVI-C | 10 | 729 | anthropogenic layer; compacted brown-grey sandy loam; inclusions of burnt clay, bones and ceramics | 0.9 | 13.9 | 2.2 | 17.9 | 4.7 |
| EVI-C1 | 11 | 734 | red-grey homogeneous feature of silty loam | 1.3 | 14.7 | 2.9 | 20.2 | 5.1 |
| EVI-D | 9 | 742 | homogeneous feature of grey ash | 1.2 | 8.4 | 14.5 | 30.9 | 4.9 |
| EVI-E | 8 | 757 | anthropogenic layer; dark brown-grey loam; inclusions of bones and ceramics | 0.2 | 19.4 | 5.5 | 15.8 | 5.1 |

of the results. Our experimental design draws on common practice during excavations: we examined selected areas of two archaeological sections of similar size (c. 1.4 x 1.0 m, Figures 6.1, 6.2) where we extracted one or two sediment samples from each stratigraphic layer, resulting in a total number of 16 samples for both profiles together. We thank the Authority for Research and Conservation of

Cultural Heritage (ARCCH) and the Tigray Culture and Tourism Bureau (TCTB) for permission to take samples and analyse them in Germany. Due to the observed differences between the two profiles analysed, regression analyses were individually performed for each profile. To ensure transparency of our results, our data is available in the supporting information (Online Resources 6.1, 6.2) along with scripts for the programming language R.

### 6.2.2.1 Laboratory analysis

We analysed 16 sediment samples, extracted from the two profiles EI and EVI. Additionally one sample of the underlying saprolite was analysed, extracted from the weathered parent material lying beneath profile EVI. Sampling locations are shown in Figure 6.2. The samples were analysed in the laboratory of Physical Geography, Freie Universität Berlin.

The water content of the sediments was determined gravimetrically, calculated according to Blume et al. (2011) and reported in mass%. The particle size distribution (1mm - 0.04$\mu$m) of the sampled material was determined with a laser diffraction particle size analyser (Beckmann-Coulter LS13 320): the prepared samples were put into a liquid sample divider and two subsamples were measured with three independent runs each. The six measurements per sample were averaged to obtain the sample's grain size distribution (Vogel et al., 2016). Particle sizes are defined according to Ad-hoc Arbeitsgruppe (2005) and reported in vol%. Analysis of element contents was conducted with (i) a Thermo Scientific Niton XL3t portable energy-dispersive X-ray fluorescence spectrometer (p-ED-XRF) and (ii) a PerkinElmer Optima 2100 DV inductively coupled plasma optical emission spectrometer (ICP-OES) after aqua regia digestion. For later statistical analyses element concentrations of Si, K, Mg and Ca measured by p-ED-XRF were used, while Na, Fe, Al, Ti and $PO_4$ concentrations were used as provided by ICP-OES data (Online Resource 6.1). Several certified reference materials (CRM) were applied for quality control: NCS DC 73325 (soil), NCS DC 73387 (soil), NCS DC 73389 (soil), LKSD-2 (lake sediment), LGC6156 (harbour sediment) and LGC6180 (flue ash). Carbon contents of the samples were examined for total carbon content (TC mass%) by burning at 1000°C in an oxygen flow (LECO TruspecCHN + S-Add-On) and inorganic carbon content (TIC mass%) by the evolution of $CO_2$ during acid ($H_3PO_4$) treatment and the subsequent quantification of the evolved $CO_2$ in 20 ml 0.05 N NaOH solution by conductivity (Woesthoff Carmhograph C-16); total organic carbon content (TOC, reported as SOM in mass%) was calculated by subtracting TIC from TC. Calcite ($CaCO_3$) contents were calculated from the measured TIC values. Mineral composition of the samples was examined with a Rigaku MiniFlex 600 X-ray powder diffractometer (XRD) with a copper $k\alpha$ tube. Mineral presence was examined semi-quantitatively using the software Philips X'Pert HighScore (v. 1.0b).

An ASD FieldSpec II spectroradiometer was used in the laboratory to capture visible (VIS) and near-infrared (NIR) reflectance of the sediment samples extracted from the profiles (350 nm - 2500 nm, 1 nm steps). Samples were measured before and after homogenisation. Each sample was illuminated by two halogen lamps. Variation of the results was minimised by capturing a white reference every 15 minutes. Each spectrum was averaged from 50 single measurements to compensate for uneven sediment texture. The captured spectra were smoothed using a Savitzky-Golay filter of second polynomial order with a width of 21 values (Savitzky and Golay, 1964; Lehnert et al., 2017). The

**Table 6.2:** Weathering indices used throughout the study. The weathering index after Parker (1970) was calculated using the atomic proportion, defined as the atomic percentage divided by atomic weight.

| Weathering Indices | Formula | Literature |
|---|---|---|
| Ruxton Ratio (RX) | $Si_2O_3 / Al_2O_3$ | Ruxton 1968 |
| Sesquioxide Ratio (SQ) | $Si_2O_3 / (Al_2O_3 + Fe_2O_3)$ | Ruxton 1968 |
| Calcium / Titanium Ratio (CTR) | $CaO / TiO_2$ | e.g. Betard 2012 |
| Potassium / Titanium Ratio (KT) | $K / Ti$ | e.g. Davies et al. 2015 |
| Parker's Index (PI) | $[(Na*/0.35) + (Mg*/0.9) + (K*/0.25) + (Ca*/0.7)] \times 100$ | Parker 1970 |

processed spectral data was used as a reference for the hyperspectral image data acquired during fieldwork.

### 6.2.2.2 Weathering indices

We used the results of the chemical sediment analyses to calculate multiple weathering indices (Table 6.2) applying ratios of certain elements or minerals to highlight differences between stratigraphic layers. Recently Zhang and Hartemink (2019c,a) showed that certain weathering indices have proven useful for the delineation of high contrast soil horizons based on image data: they observed good performance of Ca-Ti-ratio (CTR; Bétard, 2012), the Ruxton-ratio (RX; Ruxton, 1968) and the Sesquioxide-ratio (SQ; Ruxton, 1968). We also calculated these indices to test if their results can be transferred to complex archaeological profiles. Additionally we examined the K-Ti-ratio (KT) due to the high difference in ionic potential of both factors (cf. Schütt, 2004; Davies et al., 2015). Furthermore we included Parker's index (PI; Parker, 1970) as a widely accepted standard (e.g. Price and Velbel, 2001; Duzgoren-Aydin et al., 2002). Calculations of all indices include the molar mass of the respective elements.

### 6.2.2.3 Image acquisition

Image acquisition differs slightly from the setup described by Haburaj et al. (2019). Digital RGB photographs of the sections were taken using a 24.3 MP mirrorless camera with a 2.8/30 mm lens (Sony ILCE-6000, Sigma 30mm F2.8 DN Art). Hyperspectral imaging was conducted using a Cubert X2 S258 snapshot camera with a Schneider-Kreuznach Cinegon 1.8/16 mm lens. Technical specifications of the camera systems are given in Table 6.3.

The images were captured using halogen lighting (500W) to obtain uniform lighting. In total 13 hyperspectral images were recorded for profile EI and 12 hyperspectral images for profile EVI. Every single image was processed with a separate white reference measurement, allowing us to increase quality of the spectral data by including spatial variations in the lighting conditions. Additionally, each profile was captured by a single RGB image taken with the Sony ILCE-6000 camera. Black and white reference targets were used for orthorectification of RGB and hyperspectral images and the creation of overlapping data for each profile.

**Table 6.3:** Technical specifications of the camera systems used throughout this study.

|  | spatial resolution | spectral resolution | colour depth |
|---|---|---|---|
| Sony A6000 | 5696 x 4272 px | RGB in 3 bands | 8 bit |
| Cubert X2 S258 | 512 x 272 px | 450(480) - 850(830) nm in 41 (36) bands | 16 bit |

Hyperspectral images were recorded in 16-bit TIF-format using the proprietary software Cubert Utils Touch. The 8-bit images of the Sony ILCE-6000 RGB camera were also converted to TIF-format. For further processing, the pixel values of both cameras were normalised to the range [0, 1] via feature-scaling and the respective colour depth.

Since the hyperspectral images showed strong vignetting, the corners of all recordings were removed by cropping, resulting in a spatial resolution of 428 x 254 pixels. The images were stitched manually in QGIS (v3.4), using thin plate spline (TPS) transformation and nearest neighbour (NN) resampling and including the rectified RGB images as spatial reference. Histogram matching (R package RStoolbox, Leutner and Horning, 2017) was used to eliminate the remaining differences between single images. Images were then merged using GDAL (v2.4.1).

The VIS-NIR spectral data recorded with the ASD spectroradiometer for each sediment sample was used as a reference for the correction of the hyperspectral image data. A direct comparison of the two spectral datasets revealed that additional steps were necessary to eliminate noise in the image data (Figure 6.3). The merged hyperspectral image data was spatially filtered using a 5x5 median filter. The image bands from 560 to 670 nm (n=12) show high noise content since in this spectral range the measured spectrum is a composite from two overlapping sensors of the camera system (VIS sensor and NIR sensor). This range was masked and interpolated using piecewise cubic hermite interpolating polynomial interpolation (R package signal, signal developers, 2014). The spectra were then smoothed with a Savitzky-Golay filter of second polynomial order with a width of 11 values (Savitzky and Golay 1964). The image bands representing 450, 460, 470, 840 and 850 nm were excluded from the data due to their high noise content, leading to a final spectral range of 480 - 830 nm in 10 nm steps. Image cells containing missing values were interpolated using a 9x9 median filter.

A PTFE coated white reference was used for all hyperspectral recordings. The spectrum of the white reference was additionally captured using an ASD FieldSpec Handheld 2 spectroradiometer (325 - 1075 nm, 1 nm steps). The spectrum was adjusted to the spectral resolution of the hyperspectral camera and all camera recordings were divided by that spectrum to ensure accurate measurements.
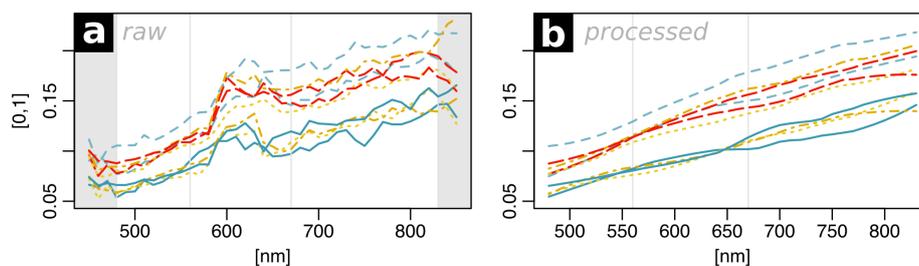


**Figure 6.3:** Spectral data before (a) and after (b) denoising. The signal was limited to the wavelength range of 480 - 830 nm and the bands between 560 and 670 nm were interpolated. The plot shows ten randomly selected spectra.

To create a uniform spatial resolution as a basis for image analyses, the RGB data was downsampled to match the resolution of the hyperspectral data. New pixel values were generated by applying a mean filter. The black and white reference targets were masked manually and remaining strong shadows were masked using threshold values for the image bands CIE $b*$ and 830 nm. This allowed us to exclude this data from the subsequent image analyses.

### 6.2.2.4 Mapping of sediment properties

Random forest regression analysis was used for the spatial mapping of sediment properties using the results of the laboratory analyses and the image data.

Various studies suggest the transformation of spectral data prior to processing (Melville and Atkinson, 1985; Viscarra Rossel et al., 2006; Stenberg and Viscarra Rossel, 2010; Hill et al., 2010; Haburaj et al., 2020a). We therefore created the following datasets as input for the regression analyses: (i) the RGB data, (ii) the CIELAB data derived from the RGB data, and (iii) the pre-processed hyperspectral data (480 - 830 nm, 10 nm steps).

Training areas for the prediction of sediment characteristics were indicated by polygons marking the sampling areas. Regression models (x,y) were trained using the spectral information from pixels (= pixel values covered by polygons) as predictors (x) and the respective sedimentological data (obtained from the laboratory analyses) as response values (y). In total we sampled c. 0.56 % of the pixels of profile EI and c. 0.28 % of the pixels of profile EVI. This led to a total of 3,581 hyperspectral and 3,455 RGB and CIELAB pixels for the seven samples of profile EI; correspondingly 2,054 hyperspectral and 1,395 RGB and CIELAB pixels for the nine samples of profile EVI were available.

Regression models were trained for profiles EI and EVI separately based on the textural, chemical and spectral characteristics of the respective sediment samples using a random forest algorithm after Breiman (2001) (R package randomForest, Cutler et al., 2018). Regression analyses were applied to SOM, $Fe_2O_3$, $CaCO_3$, the weathering indices SQ, RX, PI, KT and CTR, as well as the classified grain size (clay, silt and sand) as dependent variables and the spectral data as independent variables.

We used all sampled pixels as training data and performed a leave-p-out cross-validation with p = 0.3 over 50 iterations (R package rfUtilities Evans and Murphy, 2020). This allowed us to verify the out-of-bag errors of the random forest regressions and to report mean and standard deviation of the root mean square error (RMSE), mean bias error (MBE) and pseudo $R^2$ values $(1 - MSE/Var(y))$ from the cross-validation of each model (Cutler et al., 2018). These values were used for quality assessment of the results along with the on-site delineation of the profiles conducted by archaeological experts. The regression models were used for prediction with the image data and the R package raster (Hijmans, 2019).

## 6.2.3   Results

The deposits exposed in profiles EI and EIV correspond to a silty to sandy loose sediment, brownish to greyish in colour and with quartz and feldspar as the predominating mineral components. Individual stratigraphic features are characterised by compaction material, ash inclusions or varying amounts of clay (Table 6.1, Figures 6.2, A.1). In addition, XRD data emphasises that most samples
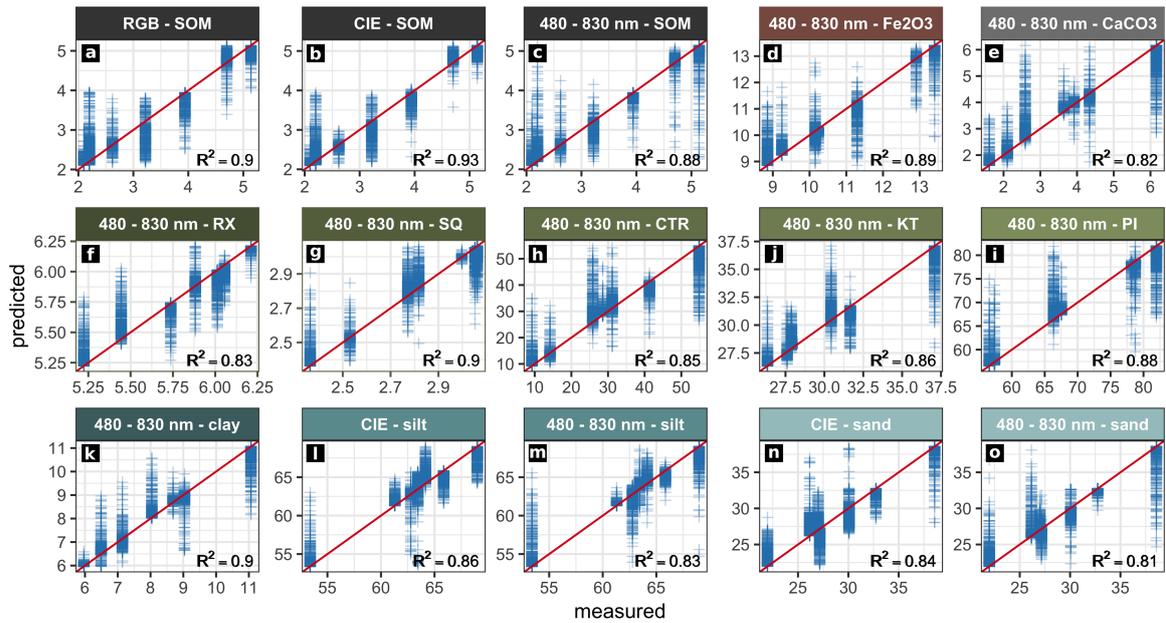
**Figure 6.4:** Results from regression analyses of profile EI. Actual (measured) and predicted values of the sediment properties examined. The ideal performance is shown by the red line (y=x). Only models with a cross-validated $R^2$ greater than 0.8 and a high correlation between the respective prediction map and the stratigraphic delineation are shown.

are characterised by distinct hematite ($Fe_2O_3$) contents (Online Resource 6.1); as Fe-bearing minerals other than hematite are circumstantial along the two sediment profiles, for the ongoing analysis the measured Fe contents are recognised as hematite components. The underlying saprolite has developed from the parent granitic bedrock and shows low carbon contents (TC: 1.55 mass%; TIC: 1.11 mass%; TOC: 0.44 mass%). Low carbonate contents all along the profiles correspond to relatively low calcium concentration, which is due to the predominance of granites in the drainage basin. Peaks of carbonate ($> 10$ mass%) only occur in the stratigraphic features EVI-B1, EVI-B2 and EVI-C1 in profile EVI. $PO_4$ concentrations along both profiles are low – slightly increased only in layers EI-A, EI-C, EI-D and EI-D1 in profile EI and showing a singular peak in layer EVI-D in profile EVI (c. 3-4.5 vol%).

The initially described differences between the two profiles are also visible in their sedimentological record: minor changes of the examined parameters are visible throughout profile EI (Table 6.1, Figure A.1), where layer EI-C shows the most distinct differences to its overlying and underlying layers in $PO_4$, SOM and $CaCO_3$ contents. Most other changes along profile EI are gradual. In contrast, in profile EVI we observed more pronounced differences between the individual layers in water content, SOM, $PO_4$, $Fe_2O_3$, $CaCO_3$ contents and the calculated weathering indices.

We trained random forest regression models between the spectral data acquired for the sampling points from the acquired image data and the examined sediment parameters (SOM, $Fe_2O_3$, $CaCO_3$, the weathering indices SQ, RX, PI, KT, CTR, classified grain size). Thereby, we were able to predict maps of these parameters which cover profiles EI and EVI. Evaluation of the trained models was carried out by leave-p-out cross-validation (Table 6.4). Only prediction maps with an $R^2$ value greater than 0.8, low values of RMSE and standard deviation and a good correlation of the prediction
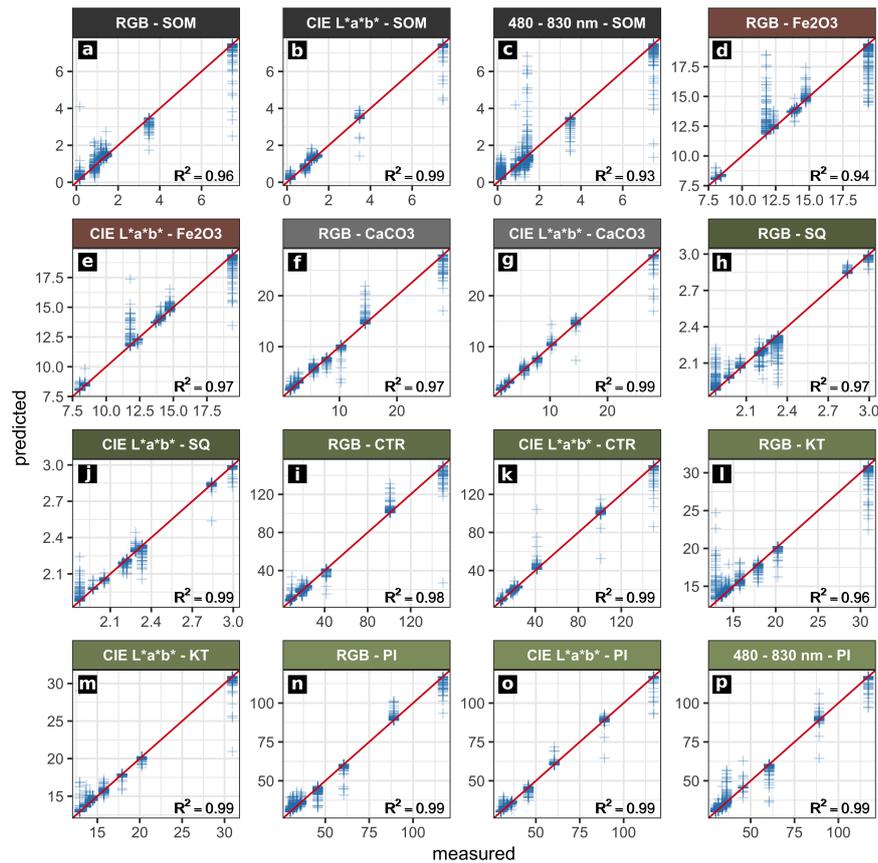
**Figure 6.5:** Results from regression analyses of profile EVI. Actual (measured) and predicted values of the sediment properties examined. The ideal performance is shown by the red line (y=x). Only models with a cross-validated $R^2$ greater than 0.8 and a high correlation between the respective prediction map and the stratigraphic delineation are shown.

results with the on-site stratigraphic interpretation are evaluated as significant and therefore depicted in Figures 6.4, 6.5, 6.6 and 6.7. To ensure transparency, all produced models and maps are available in the Online Resource 6.1.

### 6.2.3.1 Profile EI

We produced maps of multiple sediment parameters throughout profile EI by applying statistical models based on RGB and hyperspectral data. The RGB data of profile EI produced good results when predicting SOM (Figures 6.4, 6.6a, Table 6.4; $R^2$: 0.9). Similar results were obtained applying the RGB derived CIELAB data (Figure 6.6b; $R^2$: 0.93). Both sets of image data managed to roughly capture the main differences between the most prominent layers of profile EI: the increase in SOM content of layers EI-A, EI-B and EI-D. The RGB derived CIELAB data additionally captured the differences in silt and sand concentration throughout profile EI, especially visible in layer EI-E (Figures 6.6l, 6.6n). Cross-validation results for these parameters are robust with high $R^2$ and low RMSE values as well as low standard deviations (Table 6.4). These values are significantly lower for the remaining sediment parameters of profile EI. Prediction results of weathering indices, $Fe_2O_3$, $CaCO_3$ and clay, as calculated from RGB data and RGB derived CIELAB data of profile EI are therefore not reliable (Table 6.4).

**Table 6.4:** Cross-validation results of the regression analyses using RGB, CIELAB and hyperspectral data. Mean and standard deviation of the root mean square error (RMSE), mean bias error (MBE) and pseudo $R^2$ values $(1 - MSE/Var(y))$ obtained from leave-p-out cross-validation are reported. The training data consisted of seven sediment samples for profile EI and nine sediment samples for profile EVI. Significant models are marked in bold.

| | | Profile EI (n=7) | | | | | | Profile EVI (n=9) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $R^2$ | $R^2(sd)$ | RMSE | RMSE (sd) | MBE | MBE (sd) | $R^2$ | $R^2(sd)$ | RMSE | RMSE (sd) | MBE | MBE (sd) |
| SOM | RGB | **0.9** | 0.003 | 0.366 | 0.012 | 0.001 | 0.012 | **0.961** | 0.004 | 0.384 | 0.079 | -0.018 | 0.021 |
| | CIELAB | **0.932** | 0.002 | 0.301 | 0.01 | -0.001 | 0.011 | **0.987** | 0.002 | 0.214 | 0.065 | -0.005 | 0.014 |
| | 480 - 830 nm | **0.883** | 0.005 | 0.385 | 0.021 | 0.004 | 0.015 | **0.933** | 0.006 | 0.503 | 0.069 | 0.006 | 0.023 |
| $Fe_2O_3$ | RGB | 0.446 | 0.015 | 1.188 | 0.03 | 0.004 | 0.044 | **0.942** | 0.006 | 0.837 | 0.089 | -0.002 | 0.057 |
| | CIELAB | 0.599 | 0.013 | 1.016 | 0.032 | 0.002 | 0.039 | **0.974** | 0.003 | 0.556 | 0.092 | -0.001 | 0.034 |
| | 480 - 830 nm | **0.886** | 0.003 | 0.54 | 0.022 | -0.007 | 0.02 | 0.925 | 0.003 | 0.972 | 0.067 | -0.016 | 0.05 |
| $CaCO_3$ | RGB | 0.324 | 0.016 | 1.265 | 0.03 | -0.008 | 0.047 | **0.972** | 0.003 | 1.016 | 0.185 | -0.032 | 0.057 |
| | CIELAB | 0.507 | 0.015 | 1.079 | 0.032 | -0.003 | 0.038 | **0.989** | 0.002 | 0.678 | 0.195 | -0.026 | 0.043 |
| | 480 - 830 nm | **0.822** | 0.006 | 0.649 | 0.027 | 0.004 | 0.026 | 0.965 | 0.003 | 1.113 | 0.132 | -0.009 | 0.047 |
| RX | RGB | 0.53 | 0.011 | 0.21 | 0.004 | 0 | 0.008 | **0.945** | 0.004 | 0.106 | 0.011 | -0.001 | 0.007 |
| | CIELAB | 0.663 | 0.009 | 0.176 | 0.005 | 0 | 0.007 | **0.978** | 0.002 | 0.065 | 0.011 | 0.001 | 0.004 |
| | 480 - 830 nm | **0.828** | 0.004 | 0.124 | 0.004 | 0.001 | 0.005 | 0.875 | 0.01 | 0.162 | 0.017 | -0.004 | 0.008 |
| SQ | RGB | 0.457 | 0.014 | 0.18 | 0.005 | -0.001 | 0.008 | **0.973** | 0.003 | 0.058 | 0.006 | -0.002 | 0.004 |
| | CIELAB | 0.602 | 0.012 | 0.151 | 0.005 | 0 | 0.006 | **0.988** | 0.001 | 0.038 | 0.006 | 0 | 0.002 |
| | 480 - 830 nm | **0.897** | 0.003 | 0.076 | 0.003 | 0.001 | 0.003 | 0.97 | 0.001 | 0.059 | 0.004 | 0 | 0.003 |
| CTR | RGB | 0.318 | 0.015 | 12.619 | 0.298 | -0.126 | 0.499 | **0.98** | 0.004 | 5.127 | 1.69 | -0.128 | 0.297 |
| | CIELAB | 0.502 | 0.011 | 10.716 | 0.242 | -0.015 | 0.45 | **0.99** | 0.003 | 4.437 | 1.189 | -0.033 | 0.319 |
| | 480 - 830 nm | **0.847** | 0.005 | 5.938 | 0.25 | 0.066 | 0.239 | 0.99 | 0.001 | 3.777 | 0.778 | -0.058 | 0.177 |
| KT | RGB | 0.605 | 0.011 | 2.17 | 0.047 | -0.027 | 0.066 | **0.961** | 0.004 | 1.147 | 0.202 | 0.045 | 0.071 |
| | CIELAB | 0.726 | 0.008 | 1.809 | 0.045 | 0.006 | 0.066 | **0.985** | 0.003 | 0.68 | 0.217 | -0.015 | 0.045 |
| | 480 - 830 nm | **0.861** | 0.004 | 1.266 | 0.052 | 0.021 | 0.047 | 0.951 | 0.004 | 1.246 | 0.209 | -0.008 | 0.046 |
| PI | RGB | 0.478 | 0.013 | 6.9 | 0.141 | -0.035 | 0.271 | **0.987** | 0.002 | 2.782 | 0.364 | -0.102 | 0.196 |
| | CIELAB | 0.62 | 0.013 | 5.878 | 0.164 | 0.016 | 0.217 | **0.994** | 0.001 | 2.089 | 0.51 | -0.092 | 0.126 |
| | 480 - 830 nm | **0.876** | 0.003 | 3.335 | 0.156 | 0.007 | 0.115 | **0.99** | 0.001 | 2.574 | 0.435 | -0.044 | 0.145 |
| clay | RGB | 0.542 | 0.014 | 1.033 | 0.034 | 0.004 | 0.04 | 0.837 | 0.017 | 1.911 | 0.2 | -0.041 | 0.122 |
| | CIELAB | 0.661 | 0.012 | 0.897 | 0.035 | -0.006 | 0.034 | 0.929 | 0.008 | 1.271 | 0.265 | -0.041 | 0.071 |
| | 480 - 830 nm | **0.901** | 0.003 | 0.48 | 0.023 | -0.003 | 0.018 | 0.959 | 0.003 | 0.955 | 0.089 | -0.012 | 0.043 |
| silt | RGB | 0.787 | 0.008 | 1.893 | 0.058 | -0.014 | 0.057 | 0.934 | 0.004 | 3.216 | 0.337 | -0.096 | 0.182 |
| | CIELAB | **0.862** | 0.006 | 1.502 | 0.063 | 0.011 | 0.055 | 0.976 | 0.001 | 2.018 | 0.294 | -0.05 | 0.128 |
| | 480 - 830 nm | **0.834** | 0.006 | 1.665 | 0.09 | 0.014 | 0.061 | 0.898 | 0.005 | 4.009 | 0.337 | -0.038 | 0.175 |
| sand | RGB | 0.765 | 0.008 | 2.21 | 0.06 | 0.02 | 0.082 | 0.904 | 0.007 | 4.597 | 0.395 | 0.079 | 0.262 |
| | CIELAB | **0.838** | 0.006 | 1.834 | 0.072 | -0.005 | 0.076 | 0.965 | 0.003 | 2.742 | 0.444 | 0.084 | 0.21 |
| | 480 - 830 nm | **0.81** | 0.006 | 1.961 | 0.095 | 0.018 | 0.072 | 0.901 | 0.004 | 4.682 | 0.344 | 0.07 | 0.278 |
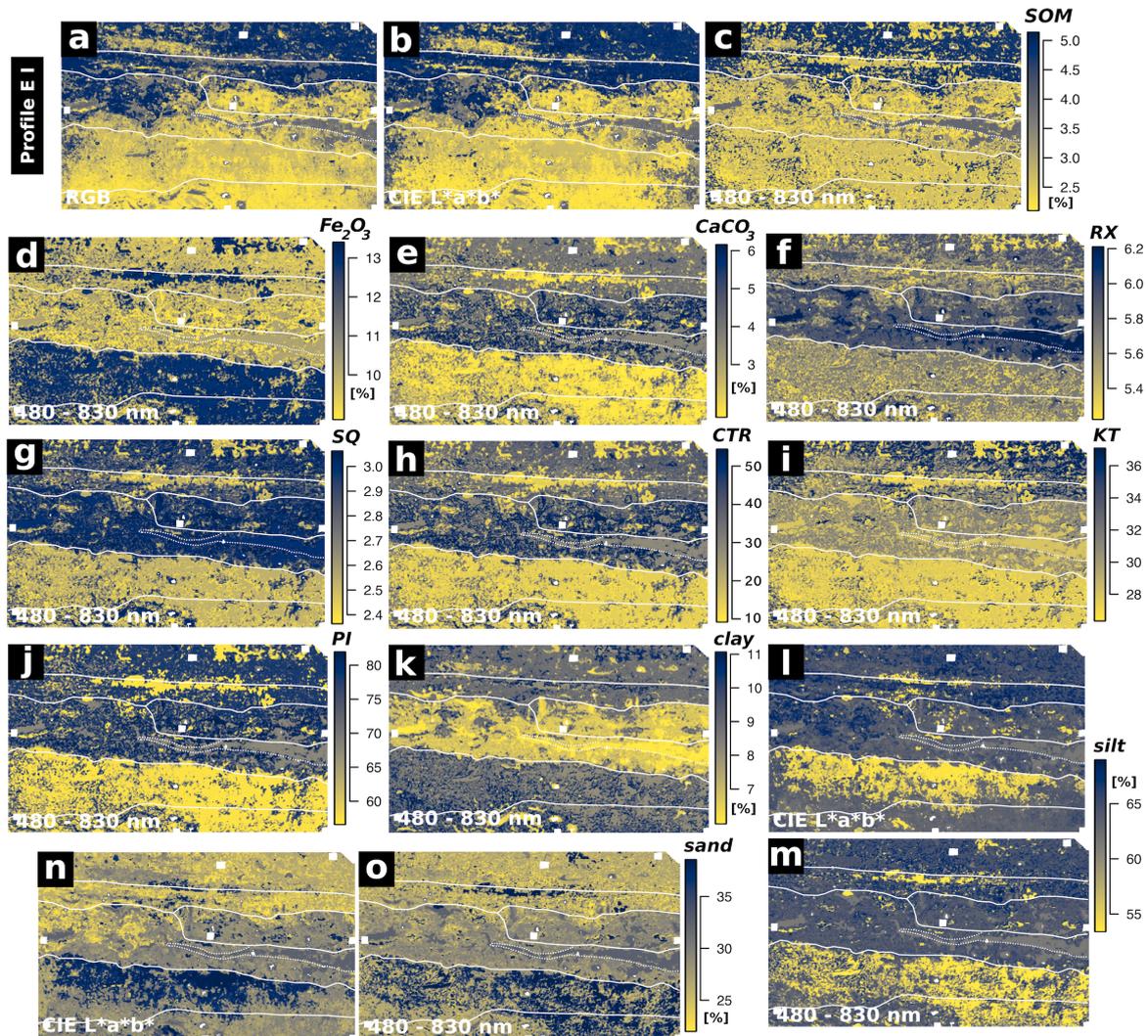
**Figure 6.6:** Predicted sediment properties of Profile EI. Only models with a cross-validated $R^2$ greater than 0.8 and a high correlation between the respective prediction map and the on-site stratigraphic delineation (white lines) are shown.

The hyperspectral image data of profile EI generally produced more accurate results for all of the predicted parameters when compared to the results obtained by applying the RGB data and the RGB derived CIELAB data. In general, we observed a good congruence between calculated and on-site delineated maps by visual comparison. Additionally, regression models calculated using the hyperspectral data of profile EI and textural and chemical sediment characters as independent variables produced $R^2$ values between 0.81 and 0.9 (n=7) and low error values (Table 6.4). These statistical results document the reliability of the parameter maps produced by these models and underline their good agreement with the on-site delineation. The most reliable models generated from the hyperspectral data of profile EI were derived for SOM contents, hematite ($Fe_2O_3$) contents, the SQ weathering index and clay contents (Figures 6.6c, 6.6d, 6.6g, 6.6k). Applying these statistical models based on the hyperspectral image data, layers EI-E and EI-F were successfully separated from the rest of the profile regarding their $Fe_2O_3$ and $CaCO_3$ contents, grain size composition and weathering indices (Figure 6.6). Additionally, layers EI-D and EI-D1 were reliably outlined by applying models of the

weathering index RX, SOM content and grain size composition based on hyperspectral image data. The predicted map of SOM successfully captures the high values present throughout layers EI-A and EI-B (Figures 6.6c, A.1). These two layers also show a low sand content in the sedimentological record (Figure A.1) and the parameter maps (Figure 6.6o).

### 6.2.3.2 Profile EVI

The random forest regression models and the resulting prediction maps produced using the RGB image data of profile EVI show good results for SOM, $Fe_2O_3$ and $CaCO_3$ contents, as well as for the weathering indices SQ, CTR, KT and PI (Figures 6.5, 6.7, Table 6.4). Increased $Fe_2O_3$ contents for layers EVI-C1 and EVI-E are clearly visible in the map based on the RGB data (Figure 6.7d). The black charcoal layer in EVI-B2 is clearly captured by the prediction map of SOM derived from the RGB data (Figure 6.7a). Prediction maps for profile EVI based on regression models calculated for the RGB derivative CIELAB are similar to those produced by RGB data (Figure 6.7). Also the $R^2$ values for both datasets are very similar for the mentioned parameters (Table 6.4; $R^2 > 0.94$, n=9). For both RGB and CIELAB data, statistical models for all weathering indices except RX outline layers EVI-B1, EVI-B2 and EVI-D (Figures 6.7h - 6.7p). Mapping results based on modelling grain size composition (sand, silt, clay) are less pronounced than observed for profile EI; statistical models involving grain size composition and the RGB data and RGB derived CIELAB data of profile EVI are not significant (Table 6.4). In contrast, the $R^2$ values of modelled grain size composition based on the hyperspectral image data are greater than 0.8. However, the resulting prediction maps for profile EVI do not match with the on-site stratigraphic interpretation (Online Resource 6.2).

Prediction maps generated by applying statistical models which involve the datasets from the hyperspectral images of profile EVI show a strong variance of the data inside the profile from left to right, indicating a systematic error; this shift was not captured by the on-site stratigraphic interpretation (Figures 6.2, 6.7). Models resulting from random forest regression analyses of the datasets from hyperspectral images and the different sedimentological parameters generally show significant $R^2$ values ($> 0.87$, n=9, Table 6.4). In contrast, parameter maps produced by applying these statistical models to the hyperspectral image data only agree in part with the maps derived from the respective statistical models based on the RGB data and the RGB and its derivatives: while for the maps derived from the application of the statistical models based on RGB data and RGB derived CIELAB data the layers outlined highly correspond with the on-site delineated stratigraphy, this does not pertain to the maps generated by the application of statistical models based on hyperspectral image data. For example, this applies to the SOM contents as shown in Figure 6.7c: SOM concentrations are generally low and layers EVI-B1 and EVI-B2 are outlined correctly with peaking SOM contents; in addition, noise in the data does not allow a reliable delineation of layers EVI-C1 and EVI-E. Apart from this SOM map of limited quality, a map of acceptable quality based on the hyperspectral data of profile EVI could only be produced for the weathering index PI (Figures 6.5p, 6.7p). All other statistical models of the hyperspectral data of profile EVI and the remaining parameters produced good statistical significances (Table 6.4) but the derived parameter maps did not hold any significant information.
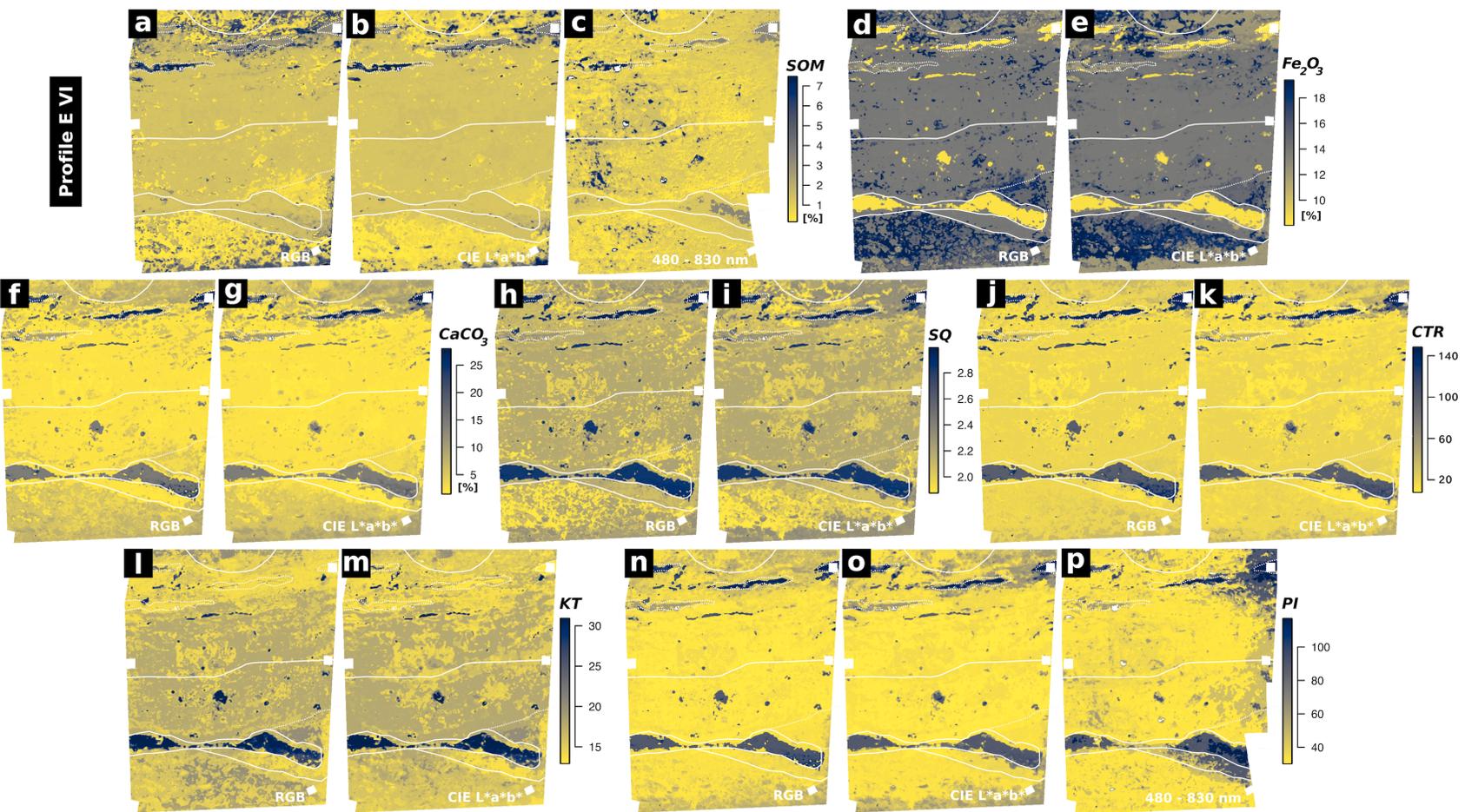
**Figure 6.7:** Predicted sediment properties of Profile EVI. Only models with a cross-validated $R^2$ greater than 0.8 and a high correlation between the respective prediction map and the on-site stratigraphic delineation (white lines) are shown.

## 6.2.4  Discussion

### 6.2.4.1 Prediction results

**General**   Modelling layer delineation for profiles EI and EVI generally shows a good performance of models based on RGB image data for predicting soil organic matter (SOM). Layer delineation by applying models based on the RGB derived CIELAB data leads to more homogeneous prediction maps of SOM contents (Figures 6.6a, 6.6b, 6.7a, 6.7b), which is consistent with the findings of Zhang and Hartemink (2019a). The good performance of the statistical models based on hyperspectral data of profile EI is consistent with the laboratory studies of, for example, Steffens and Buddenbaum (2012); Schreiner et al. (2015); Hobley et al. (2018); Heil et al. (2020) and on-site studies like Zhang and Hartemink (2019a). Applying the hyperspectral data allowed us to develop statistical models which are reliably able to map the predicted layers based on the sediment properties for profile EI. In contrast, spectral noise and disturbances of the hyperspectral data of profile EVI generated relatively unreliable results for various sediment characteristics and thus did not delineate stratigraphic layers convincingly.

**SOM**   The good results for predicting SOM from all the used image data are accounted for by the fact that the visual reflectance of soils and sediments (i.e. brightness) is highly related to SOM content (cf. Baumgardner et al., 1985). The observed increase in mapping quality of SOM when using the hyperspectral image data of profile EI as independent variable in the statistical models shows the potential of a higher spectral resolution throughout the visible part of the electromagnetic spectrum. Likewise the results are certainly influenced by the range of the hyperspectral sensor: the equidistant measurement of the area between 600 and 830 nm clearly adds important information compared to the RGB data measured with a sensor that shows rather limited capabilities beyond 650 nm (cf. Haburaj et al., 2019). However, the absolute difference between the mapping results of SOM from RGB and hyperspectral data as visible in Figures 6.6 and 6.7 is relatively low, thus stressing the potential of RGB imaging for the mapping of selected parameters, as suggested by e.g. Zhang and Hartemink (2019a).

**$Fe_2O_3$**   The iron oxide hematite shows specific spectral features between 510 and 620 nm (Baumgardner et al., 1985; Stenberg et al., 2010) and absorption bands at 450, 680 and 700 nm are also related to hematite (Baumgardner et al., 1985; Stenberg and Viscarra Rossel, 2010). The hyperspectral sensor used directly measures these wavelengths while the RGB data just covers them in one summarising red image band; in consequence, application of the hyperspectral data leads to a good performance of the statistical model when mapping $Fe_2O_3$ in profile EI. However, as the results of profile EVI reveal, RGB and RGB based CIELAB image data can clearly be used to map hematite content throughout a profile, as long as the contents are high enough. Several authors prove a direct relation between iron oxide contents and redness of soils and sediments, e.g. through the colour co-ordinates a* and b* (Barron and Torrent, 1986; Jarmer and Schütt, 1998; Vodyanitskii and Kirillova, 2016), which also explains the good performance of the random forest regression models based on RGB and RGB based CIELAB image data of profile EVI.

**Grain size**    We produced good results when mapping grain size throughout profile EI by applying statistical models based on the CIELAB and hyperspectral image data. The performance of CIELAB data was acceptable, but using hyperspectral data clearly led to more reliable maps (Figure 6.6). We see these results as indicating the potential of an extended spectral range when mapping grain size, as we were able to compensate for the lower number of samples in comparison to Zhang and Hartemink (2019a) by extending our spectral range to 480 - 830 nm. Data presented by Post and Noble (1993) and later Stenberg et al. (2010) indicates that the reliability of characterising textural elements based on statistical models increases when sensors covering NIR are included since many reflectance features related to grain size only occur beyond 1000 nm. Most studies dealing with the prediction of grain size through spectral data do indeed cover the VIS-NIR area up to 2500 nm or even use mid infrared data (MIR; 2,500 - 25,000 nm) (cf. Viscarra Rossel et al., 2006).

**Weathering indices**    Weathering index mapping based on statistical models worked out nicely in many cases and proved helpful for assessing similarities and differences between layers (see below). Statistical models based on RGB and CIELAB image data were observed by Zhang and Hartemink (2019a) to perform well in predicting the indices SQ, RX and CTR; this could be reproduced in part with our experimental design (Figures 6.6, 6.7, Table 6.4). The mapping results of the Sesquioxide ratio based on the RGB and CIELAB image data were especially helpful for delimiting layers EVI-C1 and EVI-E in profile EVI (Figure 6.7). Derived from the statistical models based on the hyperspectral image data, the prediction map of the weathering index after Parker (1970) showed a slight advantage over the weathering index CTR (derived from the same datasets) for mapping layers EI-B and EI-E / EI-F in profile EI (Figure 6.6). As the reliable prediction maps of PI and CTR are very similar in both examined profiles (Figures 6.6, 6.7), we suggest the usage of CTR over PI since less laboratory analyses are necessary to calculate the CTR index (Table 6.2). Mapping results of the index KT (K-Ti-ratio) vary from those of other indices (Figures 6.6, 6.7) and are likely related to grain size (as is also visible in Figure 6.6), as K and Ti concentration was shown to be correlated with clay, sand and silt fractions (e.g. Marshall et al., 2012).

### 6.2.4.2 Consequences of and for the stratigraphic profiles examined

Several of the examined sediment parameters were helpful to delineate stratigraphic layers and thus provide supportive information for the on-site stratigraphic description. In profile EI, the boundary between layers EI-E and EI-F (as documented during fieldwork) is not visible in the generated prediction maps, which are assessed as reliable (Figure 6.6); however, the sedimentological record (Table 6.1, Figure A.1) also only points out slight differences between both layers. Likewise the increased stone contents and the anthropogenic remains in layer EI-F (Table 6.1) are not detectable with the experimental setup used. Stone inclusions (and layer delineation based on flakes of charcoal, bones or burnt clay) in general are a crucial drawback of the proposed method, as parameter mapping results depend on the sampled materials. Circumvention of this problem would require an increased number of samples, a higher spatial resolution (smaller pixel size) and an additional layer of documentation in the form of image texture or similar parameters (cf. Haburaj et al., 2019). However, based in the fines

of layers EI-E and EI-F (Profile EI), both strata are delineated nicely from the rest of the profile in the parameter maps (Figure 6.6). Layers EI-A and EI-B (Profile EI) are mainly distinguished from each other by their differing density of inclusions (Table 6.1), again causing a relatively weak delineation between the two strata based on the application of statistical models applying RGB, its derivative CIELAB or hyperspectral data as independent variables (Figure 6.6). The on-site description of layer EI-B as 'more homogeneous' (Table 6.1) compared to layer EI-A, however, is underlined by a horizontal patch of homogeneous material visible in the mapping results of $Fe_2O_3$, $CaCO_3$ and PI, as derived from the hyperspectral image data by statistical modelling (Figure 6.6). The generally high contrast between layers EI-A and EI-B regarding weathering indices (Figure 6.6) supports the on-site description of the two layers as individual strata.

In Profile EVI, the examined sediment properties also provide supportive information for the on-site stratigraphic description. The parameter maps produced by applying the reliable statistical models show similarities between layers EVI-C1 and EVI-E (Figure A.1, 7). Increased $Fe_2O_3$ contents and similar values of the SQ weathering index are visible for both layers in the respective parameter maps (Figure 6.7) and the increased iron content explains the red-brown colour of the two layers EVI-E and EVI-C1 (cf. Hurst, 1977, Table 6.1). The on-site description of layers EVI-D and EVI-B1 as similar material (bright ash layers) is not visible in our results. Rather, we observed clear differences between the two layers: the parameter maps of the weathering index KT produced by applying the statistical models including the RGB and CIELAB image data, indicate a significantly different weathering grade of layers EVI-D and EVI-B1 (Figure 6.7). This is also visible in the calculated KT and RX values in the sedimentological record of profile EVI (Figure A.1). Additionally the two layers clearly differ in $PO_4$ concentration (Figure A.1). Layer EVI-B1 can clearly be described as a bright ash layer connected with anthropogenic activities. In contrast, we interpret layer EVI-D as a thin layer covering layers EVI-E and EVI-C1 and most likely related to a reshaping of the surface or construction work. Remains of this process are also visible on the opposite side of the trench at the same depth (Japp, 2019). The lowermost anthropogenic layer (layer EVI-E) varies greatly in thickness and locally shows a gradual transition into the underlying saprolite (cf. Stolt and Baker, 1994; Japp, 2019).

### 6.2.4.3 Reproducibility

Evaluating our results, clear discrepancies are visible between profiles EI and EVI regarding the performance of the RGB, RGB derived CIELAB and hyperspectral data. For profile EI, using statistical models which are based on the hyperspectral data led to a general increase of the quality of the prediction results when compared to the results achieved by the RGB and the derived CIELAB data (Figures 6.4, 6.6, Table 6.4). In contrast, using the hyperspectral data of profile EVI in most cases led to more inaccurate parameter maps than the maps produced on the basis of the RGB and the derived CIELAB data. Together with the irregular cross-validation results of the hyperspectral data of profile EVI (Table 6.4), we interpret this as resulting from the observed variations, noise and disturbances in the hyperspectral image data of profile EVI, presumably caused by light that was irregularly reflected by the adjacent profile walls between the recordings of the 12 hyperspectral images of profile EVI.

The two profiles additionally differ significantly in their overall characters: profile EI is a low

contrast profile and EVI is a high contrast profile. Pronounced differences in lightness and redness are only visible in profile EVI (Figure 6.2). These characteristics result from a high variation of SOM contents (lightness) and $Fe_2O_3$ contents (redness) throughout the profile. As lightness and redness heavily influence the spectral reflectance in the visual part of the electromagnetic spectrum, the statistical models based on the RGB data show a better performance for profile EVI than for the low contrast profile EI (Figures 6.6, 6.7).

Prediction results of both profiles are influenced by our sampling strategy, which we limited to less than ten sediment samples per profile. The digitised training areas for the regression analyses were focused on the direct surroundings of the samples and thus only covered a small percentage of the total image pixels ($< 0.6$ %). This low number of sediment samples leads to a shorter processing time, and thus lower costs, compared to other studies (e.g. Hobley et al., 2018; Zhang and Hartemink, 2019a). Generally, the gradual and rather low variation of the parameters examined throughout profile EI was captured successfully despite the low amount of samples (n = 7). Profile EVI, however, shows a more complex stratigraphy that probably would have required a more detailed sampling strategy to be able to outline its stratigraphic character: (i) singular lenses of very dark and very bright material, (ii) gradual variation in vertical and horizontal directions, and (iii) many anthropogenic layers which show differing contents of bones, charcoal, ceramics and burnt clay. These characteristics may lead to a disproportional representation of some of the examined sediment properties in the samples and thus in the parameter maps derived from the statistical models based on these samples. Nonetheless, we see the main reason for the poor performance of the hyperspectral data of EVI in the lighting situation during image acquisition: the adjacent walls of the trench are situated directly left and right of the captured part of profile EVI. Similar to the shadow cast by adjacent trench walls as described by Haburaj et al. (2020a), the walls in profile EVI most likely reflected and scattered the artificial light differently between single measurements due to changing lamp positions and caused the observed distortions in the hyperspectral image data of profile EVI.

Our results clearly show that a low number of samples can be sufficient to successfully map parameters like the concentrations of organic matter or iron oxide. Nevertheless, the samples need to be representative for the profile examined. Complex anthropogenic layers featuring high amounts of flakes of brick, charcoal, bones or mortar can be reliably delineated with a very high spatial resolution of the image data and the sampling strategy. Apart from the contrast of the examined profile or planum, the spatial scale and structure should be kept in mind. Similar problems and thoughts have already been discussed in the field of remote sensing: additionally to the sizes and spatial relationships of the objects in the recorded scene (as suggested by Woodcock and Strahler, 1987), the spectral relationships of these objects should be considered when assessing the required scale of a scene.

As the CIELAB image data used throughout the presented study is derived from the acquired RGB image data, it should be kept in mind that the parameter maps of the two datasets are highly correlated (e.g. Pearson correlation coefficient of 0.92 with p $< 0.05$ between SOM values predicted from RGB and CIELAB data, see Online Resource 6.2). Transformation of the RGB data to the CIELAB colour space, however, has some benefits: the parameter maps based on the CIELAB data frequently showed less noise (Figures 6.6, 6.7). Results of the cross-validations also indicate a slightly better performance of the statistical models based on the CIELAB data when compared to the models
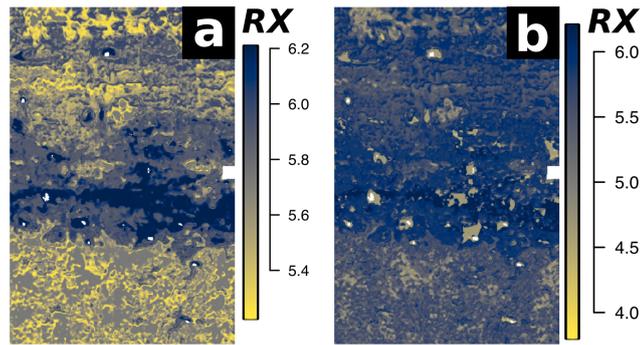
**Figure 6.8:** Predicted weathering index values (RX = Ruxton Ratio) of Profile EI. Prediction results from the regression models trained with (a) the hyperspectral image data of profile EI and (b) the combined hyperspectral image data of profiles EI and EVI. When analysing the profiles individually, significantly more detail is preserved (a).

directly based on the RGB data (Table 6.1). These findings are consistent with multiple studies which suggest that the CIELAB colour space shows clear benefits for quantitative analyses (e.g. Melville and Atkinson, 1985; Vodyanitskii and Kirillova, 2016).

In the presented study we produced significantly better mapping results when training regression models for profiles EI and EVI individually. When applying the random forest regression analysis to the merged image data and sedimentological data of both profiles, much detail was removed from the resulting prediction maps (Figure 6.8). We assume that the high contrast of profile EVI caused by just a few sediment properties resulted in these observed differences in modelling performance.

### 6.2.4.4 Accuracy

Overall accuracy of the parameter maps derived from the statistical models was assessed (i) by visual comparison to the on-site stratigraphic interpretation (Figure 6.2), (ii) comparison of the mapping results to the laboratory results describing stratigraphic layers (Table 6.1, Figure A.1) and (iii) by performing leave-p-out cross-validation (CV), resulting in RMSE, MBE and pseudo $R^2$ values for each statistical model (Table 6.4). CV results of profile EI mostly agree with the quality of the mapping results of the parameters examined. They therefore help to quantify differences in the performance of the statistical models based on the image datasets used and likewise strengthen the significance of our results in general. Since our sampling strategy only allowed for the definition of highly local training and testing data, the CV results are influenced by spatial autocorrelation (Legendre, 1993). This becomes visible especially in profile EVI, where $R^2$ values are generally too high (Table 6.4). The selective character of the training data of the statistical models in combination with the distorted image data leads to an overfitting of the random forest model and thus produces the observed CV results. This is visible for multiple parameters in profile EVI: the high $R^2$ values for grain size, for example, disagree with the actual parameter maps, which are partly very inaccurate (Online resource 6.2). Another example is the disproportionally high $R^2$ value of 0.99 for the prediction of many parameters via the hyperspectral image data of profile EVI (Table 6.4). Possible solutions for this problem include tuning of the random forest model (Segal, 2004) or the usage of a spectral sensor that produces cleaner spectra and allows one to use a more specific or an extended feature space for

87

training the statistical models used for prediction (e.g. Viscarra Rossel et al., 2006). Additionally our approach only included 0.56 % (EI) and 0.28 % (EVI) of the image pixels as training data. Increasing this number may lead to more reliable prediction results (Millard and Richardson, 2015). Taking into account the reliable results for profile EI and the acceptable results produced using the RGB or RGB derived CIELAB data for profile EVI, we suggest that the poorer performance of the proposed method for profile EVI originates in the distorted hyperspectral image data. For many of the parameters examined we were able to produce good prediction results comparable with those of past studies conducted under controlled conditions in the laboratory (e.g. Hobley et al., 2018) or with a significantly higher number of samples (e.g. Zhang and Hartemink, 2019a). The expressiveness of the pseudo $R^2$ values used throughout the present study remains limited, but, in combination with the on-site layer delineation, the values led to a more traceable stratigraphic interpretation of the profiles examined.

## 6.2.5   Conclusions

This study documents that the supervised regression of sediment parameters and – RGB and hyperspectral – image data provides valuable support for on-site sediment profile delineation and characterisation by an expert (here: archaeologist). The expert knowledge applied during on-site delineation of a sediment profile allows an integral analysis including the detailed description of certain macroscopic aspects such as texture, colour or compactness. Supervised regression involving image data is exclusively based on the physical spectral signal and requires the generation and application of several statistical models, mostly one per parameter contemplated. However, by applying these models, various chemical and textural sediment characteristics can be delineated across a profile and thus add information to the expert's delineation and interpretation of layers.

We demonstrate that the extrapolation of sediment properties of sampled material via image data can be transferred from the laboratory to archaeological fieldwork conditions. We were able to create reliable maps of multiple sediment parameters (SOM, $Fe_2O_3$, $CaCO_3$, weathering, grain size) by applying statistical models for each parameter. The low number of samples furthermore renders our approach very limited in its destructiveness, which is crucial when researching anthropogenic heritage.

As the results imply, several characteristics of the proposed method must be considered: (i) the fieldwork situation needs to be carefully controlled, (ii) artificial light is highly recommended, and (iii) the spatial resolution of the image data and the sampled materials has to be adjusted according to the size, scale and complexity of the respective profile or planum examined. Bearing this in mind, our results suggest that spectral sensors in general, and particularly more advanced spectral sensors covering the NIR and MIR region in more detail (Linderholm et al., 2019), are suitable for the mapping of multiple sediment properties, commonly examined selectively during excavations. Our study illustrates the potential held by digital imaging as applied as a standard feature during most excavations nowadays. Treating digital images as physical measurements clearly offers a quick and low-cost way to increase the traceability of stratigraphic delineation and interpretations.

Future studies could also apply the proposed method to archaeological plana, possibly recorded by a multi- or hyperspectral sensor mounted on an unmanned aerial vehicle. However, as spectral information should serve to support profile delineation and not substitute it, we suggest a combined application of expert knowledge and spectral data.

## Acknowledgments

# CHAPTER 7

## Conclusions and Outlook

## 7.1 Major conclusions

**Spectroscopy as a complementary method**   The presented results clearly show that quantitative colour measurements and spectral measurements acquired during fieldwork can complement traditional conventions of stratigraphic documentation and interpretation. We suggest the semi-automatic delineation of stratigraphic layers, based on a profound statistical analysis of quantitative data, as an opportunity to verify and support existing ways of delimiting diagnostic horizons. Following the path laid out by Spinner (1974), the proposed methodologies offer a gain of knowledge by increasing accuracy, certainty, traceability and reproducibility.

It is shown that the identification of stratigraphic layers can be carried out in a traceable way by applying unsupervised machine learning algorithms. Specifically, the unsupervised classification of image data offers additional information that may help researchers draw and characterise stratigraphic borders and quantify the homogeneity of the delineated layers (cf. chapter 4, Haburaj et al., 2019). By treating digital image data as physical measurements, the methodology proposed in this thesis offers a solution for two significant problems: (i) the destructive character of excavations demands documentation that gathers as much information as possible, and (ii) the common ways of documentation are often influenced by the subjective perception of the researchers at work. Besides, the results presented in chapter 5 show that unsupervised classification is also a promising tool when dealing with selectively sampled data (cf. Haburaj et al., 2020a). Clustering of selectively sampled hyperspectral data produces results of similar quality as the clustering of p-ED-XRF data, when defining the stratigraphic sequence. Furthermore, since such spectral measurements are not relying on destructive laboratory work, but rather constitute a fast measurement technique that often can be applied on-site, the potential of optical data for a traceable documentation during fieldwork is further underlined. Depending on the desired output, implementation of RGB imaging, multispectral

imaging, hyperspectral imaging or selective hyperspectral readings contributes significantly to this matter.

A major advantage of digital image data over sedimentological data is its extensive character as the latter is based on selectively sampled material; i.e., by recording only a few images, one can capture entire profiles or plana. When including a systematic way of image acquisition, this trait of image data leads to a vast pool of quantitative data, which can be utilised for further analyses. Furthermore, statistical models based on such image data can be used to create reliable maps of various chemical and physical sediment properties along profiles (cf. Haburaj et al., 2020b). These maps offer extensive information which complements stratigraphic documentation, and likewise aloow the respective stratigraphic interpretation to be more transparent. Sediment properties which show a strong influence on the spectral reflectance in the visible part of the electromagnetic spectrum (e.g. soil organic matter, iron oxides) can be assessed using conventional RGB cameras, while more complex parameters (e.g. grain size) can be assessed more reliably using a hyperspectral sensor (cf. Haburaj et al., 2020b).

Overall, the results of this thesis indicate a high potential of RGB imaging for quantitative analyses. The delineation of stratigraphic layers based on unsupervised classification and the creation of parameter maps of high-contrast sediment profiles were successfully carried out using RGB data. The results presented here (cf. chapters 4, 5, 6) are consistent with other contemporary studies, adding this approach as part of a current trend towards the exploitation of RGB data as quantitative data. For example, the results of Zhang and Hartemink (2019a) show that the stratigraphy of high-contrast soil profiles could be captured by analysing RGB data, whilst Heil et al. (2020) utilise digital RGB photography to analyse various parameters of soils on a microscale. In most cases, RGB imaging devices are part of the standard equipment brought to archaeological or geoscientific fieldwork anyhow (e.g. Sapirstein and Murray, 2017). We therefore want to underline the observed additional value of such data.

The methodology provided in this thesis, however, does not constitute a ready-to-use approach. Rather, this thesis presents a *possible* new direction for the quantitative analysis of digital image data. While the methodology should therefore be evaluated as experimental, the provided results expose a reliable performance of the used statistical models (Haburaj et al., 2019, 2020a,b). Further development of the methodology is facilitated by the open-source scripts available in the appendix. The question remains, however, whether a ready-to-use approach *can* be offered at this point: as shown in this thesis, statistical modelling needs to be adapted to site-specific conditions (cf. chapter 4, Haburaj et al., 2019) as even intra-site differences could be too high to use a single statistical model for multiple profiles or plana (cf. chapter 6, Haburaj et al., 2020b). Additionally, the spectral devices, which are currently available, differ significantly in their spectral resolution and range, as well as their spatial resolution and acquisition technique (see chapter 2). Apart from the latter, these characteristics also differ between RGB cameras of different manufacturers. Therefore, despite showing highly effective results, further development and standardisation of the current workflow are highly recommended.

**Recommendations for future adaptions**   Due to the experimental character of this thesis, some recommendations for future adaptions of the proposed methodology can be formulated based on the results of the presented studies.

First, when working with optical sensors in general, one has to heavily control the fieldwork situation. Depending on the sensor used, small differences in lighting can have a significant impact on the consistency of the data. We suggest, either using artificial light sources during fieldwork (preferably after nightfall, see chapter 6, Haburaj et al., 2020b), or to work during the day, capturing the respective scene illuminated by diffuse daylight (see chapter 5, Haburaj et al., 2020a). Regardless of the imaging sensor used, special attention should be given to consistent measurements of an accurate white reference. A robust standard white reference offers consistency and comparability between single recordings and therefore should also be regarded useful for common RGB photography as conducted during excavations. As shown in chapter 2, a PTFE coated cloth can easily be manufactured into a custom-sized and highly accurate white reference.

Secondly, when mapping sediment parameters based on statistical models, which involve spectral data and chemical or physical sedimentological data (see chapter 6, Haburaj et al., 2020b), special attention should be given to the sampling and image acquisition strategy. Together with the spatial resolution of the image data and the complexity and contrast of the profile, the number of reference samples has a significant impact on the resulting parameter maps. As discussed in chapter 6, a high or low number of samples can be more effective, depending on the examined scene and the respective research question.

The third issue addressed for future adaptions concerns the hardware. Spectral cameras used throughout this thesis were chosen, as they offer a high spectral and spatial resolution compared to similar devices and are lightweight and energy-efficient (see chapter 2). Therefore they include many characteristics, which are necessary for fieldwork in remote places. As shown in chapter 6, especially the hyperspectral sensor used, however, requires a highly controlled on-site situation, including accurate artificial lighting. Additionally, the spatial resolution of the hyperspectral sensor is high, when compared to other snapshot-style cameras, but still considerably low when recording larger areas (cf. chapter 6, Haburaj et al., 2020b). Significant effort during fieldwork was necessary to produce high-quality spectral images with the used hyperspectral camera. The orthophotos used in chapter 6 were the result of multiple recordings, which had to be combined and adjusted manually. Taking this into account, the usage of a hyperspectral scanner could offer some benefits for future fieldwork studies: a higher quality of the spectral recordings, better spatial resolution and a larger area which can be captured at once. Furthermore, the increased weight and power consumption of a hyperspectral scanner become irrelevant in many cases, as necessary artificial lighting in the form of halogen lamps requires constant energy supply. However, the utility of snapshot cameras, as used throughout this thesis, should not be underestimated: (i) on-site conditions can limit the available space for setting up devices (cf. chapter 6, Haburaj et al., 2020b), (ii) the power supply can be very limited in remote places, and (iii) the spectral range and resolution of the used sensor should depend on the respective research question. The continuous technological advancements in the field of spectral imaging suggest that, in the future, spectral sensors become available, which offer the high quality spectral data of hyperspectral scanners and the light weight of snapshot-style devices. However, at the current point,

researchers have to choose a spectral camera or scanner according to the respective application based on the mentioned device specifications and fieldwork conditions.

## 7.2  Outlook

The conducted statistical modelling of sediment properties based on image data produced reliable parameter maps for two archaeological profiles (cf. chapter 6, Haburaj et al., 2020b). On the other hand, the proposed worklfow still includes time- and cost-intensive laboratory analyses of sediment samples. This laboratory based data is used to train the statistical models used for the creation of the parameter maps. By a complete integration of spectral measurements into current workflows, one could produce regional or site-specific databases which include quantitative data regarding sediment properties and spectral data: each sample which is analysed for its chemical or physical properties could additionally be measured with a spectroscopic device. A similar approach was, for example, conducted on a global scale by Viscarra Rossel et al. (2016). Such databases could offer a reliable basis for the extrapolation of the examined sediment properties based on image data. The usage of statistical models based on hyperspectral data for predictions based on optical data of lower quality was successfully applied in several studies Hill and Schütt (e.g. 2000). However, due to the sensitivity of the prediction process to changes in training data (cf. chapter 6, Haburaj et al., 2020b), the proposed databases should only include local and regional data. When successfully established, such a database would allow for more automation in the proposed workflow.

The methodology included in this thesis offers a basis, from which statistical models, which contribute to a more traceable documentation, can be further improved. The delineation of stratigraphic layers using unsupervised classification of image data (cf. chapter 4, Haburaj et al., 2019) could be further enhanced by an in-depth analysis of various classification algorithms. One recent study, for example, suggested the usage of a fuzzy approach for layer delineation (Zhang and Hartemink, 2019c). Adopting such fuzzy classification to the proposed workflow would make the layer delineation even more transparent, as the membership of pixels to certain stratigraphic layers could be determined.

The creation of parameter maps based on image data and statistical models could benefit from regional or site-specific databases (see above). Advanced machine learning algorithms could be trained using such a database and potentially produce more accurate results as presented in chapter 6 (Haburaj et al., 2020b). A promising approach was recently explored by Ng et al. (2019), who used convolutional neural networks (CNN) for the prediction of soil properties based on spectral data. CNN models show a promising performance, and can be analysed for the criteria which underlie the decisions made by the algorithm – i.e., spectral wavelengths, which are relevant for specific soil or sediment properties, can be identified (cf. Ng et al., 2019). Such an approach would help with the semi-transparency of other algorithms (e.g. random forests, neural networks, support vector machines, cf. Dinov, 2018). This is even more relevant in the context of stratigraphic analysis, where we want to utilise quantitative analyses to compensate for the black box character of traditional methods.

Replacement of one black box system (based on human perception) with another black box system (based on image data) is not a desirable outcome.

Moreover, the influence of chemical and physical characteristics on the spectral reflectance of soils and sediments could be further exploited. Hill and Schütt (2000) showed that soil organic matter can be mapped reliably by a parameterisation of the spectral signal. Models which would include the shape of the reflectance spectrum (e.g. Hill and Schütt, 2000) or specific peaks or absorption bands of the spectrum (e.g. Stoner and Baumgardner, 1981; Post and Noble, 1993; Viscarra Rossel and Behrens, 2010; Soriano-Disla et al., 2014) could clearly lead to more accurate results when mapping chemical and physical properties. The training of such a model would, however, require a spectroscopic device of high spectral accuracy like, for example, a hyperspectral scanner or a high resolution spectroradiometer (cf. chapter 5, Haburaj et al., 2020a).

Furthermore, the findings of this thesis could be combined by performing an unsupervised classification based on the resulting parameter maps. These may include maps of organic matter, chemical weathering or iron oxides, providing a larger basis for the delineation of stratigraphic layers than that used in chapter 4 (Haburaj et al., 2019). Promising results addressing this issue were achieved by Steffens and Buddenbaum (2013) and more recently Zhang and Hartemink (2019a). The results presented in chapter 6 (Haburaj et al., 2020b) suggest that this methodology can be transferred to regular fieldwork conditions in the future.

With the continuing technological progress, spectral cameras become more accurate, lightweight and offer a higher spectral and spatial resolution. In particular, higher spatial resolution of spectral recordings should be examined in the future as it could help to grasp small debris of charcoal, bones, brick or mortar included in soil and sediment layers. As indicators of human activity, these features often form the basis for the delineation of stratigraphic layers.

Extending the approach presented in this thesis, snapshot-style spectral cameras could also be attached to unmanned aerial vehicles to capture larger areas like excavated surfaces during rescue excavations related to pipeline construction. The semi-automatic analysis of this data could, for example, allow for delineating anthropogenic features like post-holes or ditches. As shown in chapter 6, many features related to human activity could be captured via their chemical or physical composition using spectral data (e.g. areas for iron smelting, settlements, stables, latrines). Depending on the examined materials, the results presented in this thesis indicate that even airborne RGB imaging could be used for this purpose.

Overall, the studies included in this thesis show that an implementation of spectral recordings and quantitative colour measurements into current workflows for stratigraphic documentation and interpretation offers clear benefits. The delineation of diagnostic layers based on image data (Haburaj et al., 2019) or selectively sampled spectral data (Haburaj et al., 2020a) adds traceability to current practices. The extrapolation of soil and sediment characteristics via image data (Haburaj et al., 2020b) further strengthens this transparency and reproducibility, and likewise offers new ways of looking at exposures of archaeological and geoscientific excavations. Future technological improvements in the field of spectral imaging, will undoubtably improve the production of reliable results based on data that was acquired qith a common RGB camera.

# Bibliography

Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., and Süsstrunk, S. (2012). SLIC Superpixels Compared to State-of-the-Art Superpixel Methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(11):2274–2282.

Ad-hoc Arbeitsgruppe Boden (2005). *Bodenkundliche Kartieranleitung (KA5)*. Hrsg. von der Bundesanstalt für Geowissenschaften und Rohstoffe in Zusammenarbeit mit den Staatlichen Geologischen Diensten, Hannover, 5 edition.

Adão, T., Hruška, J., Pádua, L., Bessa, J., Peres, E., Morais, R., and Sousa, J. J. (2017). Hyperspectral Imaging: A Review on UAV-Based Sensors, Data Processing and Applications for Agriculture and Forestry. *Remote Sensing*, 9(11):1110.

Aitchison, J. (1986). *The Statistical Analysis of Compositional Data*. Monographs on Statistics and Applied Probability. Chapman & Hall Ltd., London.

Altinbas, U., Kurucu, Y., Bolca, M., and El-Nahry, A. (2005). Using advanced spectral analyses techniques as possible means of identifying clay minerals. *Turkish Journal of Agriculture and Forestry*, 29(1):19–28.

Anfray, F. (1967). Matara. In *Annales d'Éthiopie*, volume 7, pages 33–88. Editions de la Table Ronde.

Aqdus, S. A., Hanson, W. S., and Drummond, J. (2012). The potential of hyperspectral and multi-spectral imagery to enhance archaeological cropmark detection: a comparative study. *Journal of Archaeological Science*, 39(7):1915–1924.

Bareth, G., Aasen, H., Bendig, J., Gnyp, M. L., Bolten, A., Jung, A., Michels, R., and Soukkamäki, J. (2015). Low-weight and UAV-based Hyperspectral Full-frame Cameras for Monitoring Crops: Spectral Comparison with Portable Spectroradiometer Measurements. *Photogrammetrie - Fernerkundung - Geoinformation*, 2015(1):69–79.

Barnes, R. J., Dhanoa, M. S., and Lister, S. J. (1989). Standard Normal Variate Transformation and De-trending of Near-Infrared Diffuse Reflectance Spectra. *Applied Spectroscopy*, 43(5):772–777.

Barron, V. and Torrent, J. (1986). Use of the Kubelka—Munk theory to study the influence of iron oxides on soil colour. *Journal of Soil Science*, 37(4):499–510.

Baumgardner, M., Silva, L., Biehl, L., and Stoner, E. (1985). Reflectance properties of soils. *Advances in Agronomy*, 38:1–44.

Bebermeier, W., Hoelzmann, P., Meyer, M., Schimpf, S., and Schütt, B. (2018). Lateglacial to Late Holocene landscape history derived from floodplain sediments in context to prehistoric settlement sites of the southern foreland of the Harz Mountains, Germany. *Quaternary International*, 463:74–90.

Behmann, J., Acebron, K., Emin, D., Bennertz, S., Matsubara, S., Thomas, S., Bohnenkamp, D., Kuska, M. T., Jussila, J., Salo, H., Mahlein, A.-K., and Rascher, U. (2018). Specim IQ: Evaluation of a New, Miniaturized Handheld Hyperspectral Camera and Its Application for Plant Phenotyping and Disease Detection. *Sensors*, 18(2):441.

Belušič, G., Ilić, M., Meglič, A., and Pirih, P. (2016). A fast multispectral light synthesiser based on LEDs and a diffraction grating. *Scientific Reports*, 6:32012.

Bent, T. (1893). *The sacred city of the Ethiopians*. Longmans & Green, London.

Bergh, M. V. d., Boix, X., Roig, G., and Gool, L. V. (2015). SEEDS: Superpixels Extracted Via Energy-Driven Sampling. *International Journal of Computer Vision*, 111(3):298–314.

Bernatchez, J. A. and Marean, C. W. (2011). Total station archaeology and the use of digital photography. *The SAA Archaeological Record*, 11(3):16–21.

Bishop, J. L., Pieters, C. M., and Edwards, J. O. (1994). Infrared Spectroscopic Analyses on the Nature of Water in Montmorillonite. *Clays and Clay Minerals*, 42(6):702–716.

Bivand, R. S., Pebesma, E., and Gomez-Rubio, V. (2013). *Applied spatial data analysis with R, Second edition*. Springer, New York.

Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(1):2–16.

Blume, H.-P., Stahr, K., and Leinweber, P. (2011). *Bodenkundliches Praktikum: Eine Einführung in pedologisches Arbeiten für Ökologen, Land- und Forstwirte, Geo- und Umweltwissenschaftler*. Springer, Heidelberg, 3 edition.

Boogaart, K. G. v. d., Tolosana-Delgado, R., and MatevzBren (2018). compositions: Compositional Data Analysis. https://CRAN.R-project.org/package=compositions.

Brandt, S. A. (1984). New perspectives on the origins of food production in Ethiopia. In *J. D. Clark & S. A. Brandt (Eds.), From Hunters to Farmers: the causes and consequences of food production in Africa*, pages 173–190. University of California Press, Berkeley.

Bratisi, M., Liritzis, I., Vafiadou, A., Xanthopoulou, V., Palamara, E., Iliopoulos, I., and Zacharias, N. (2018). Critical assessment of chromatic index in archaeological ceramics by Munsell and RGB: novel contribution to characterization and provenance studies. *Mediterranean Archaeology and Archaeometry*, 18(2):175–212.

Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1):5–32.

Brenning, A. (2012). Spatial cross-validation and bootstrap for the assessment of prediction rules in remote sensing: The R package sperrorest. In *2012 IEEE International Geoscience and Remote Sensing Symposium*, pages 5372–5375.

Brown, D. J., Shepherd, K. D., Walsh, M. G., Dewayne Mays, M., and Reinsch, T. G. (2006). Global soil characterization with VNIR diffuse reflectance spectroscopy. *Geoderma*, 132(3):273–290.

Brunke, H., Bukowiecki, E., Cancik-Kirschbaum, E., Eichmann, R., Ess, M. v., Gass, A., Gussone, M., Hageneuer, S., Hansen, S., Kogge, W., May, J., Parzinger, H., Pedersén, O., Sack, D., Schopper, F., Wulf-Rheidt, U., and Ziemssen, H. (2016). Thinking Big. Research in Monumental Constructions in Antiquity. *Space and Knowledge*, pages 250–305.

Bryan, P. G., Corner, I., and Stevens, D. (2003). Digital Rectification Techniques for Architectural and Archaeological Presentation. *The Photogrammetric Record*, 16(93):399–415.

Buyssens, P., Gardin, I., and Ruan, S. (2014). Eikonal based region growing for superpixels generation: Application to semi-supervised real time organ segmentation in CT images. *IRBM*, 35(1):20–26.

Bétard, F. (2012). Spatial variations of soil weathering processes in a tropical mountain environment: The Baturité massif and its piedmont (Ceará, NE Brazil). *CATENA*, 93:18–28.

Böttcher, K. (2002). Exkursionsführer: Bodenlehrpfad Chorin-Weinberg. *Annex to: Böttcher, K.: Konzeptentwicklung und Charakterisierung ausgewählter Böden für den Bodenlehrpfad Chorin-Weinberg. Eberswalde, Fachhochschule, Diploma Thesis.*

Cai, W., Chen, S., and Zhang, D. (2007). Fast and robust fuzzy c-means clustering algorithms incorporating local information for image segmentation. *Pattern Recognition*, 40(3):825–838.

Chen, S., Li, S., Ma, W., Ji, W., Xu, D., Shi, Z., and Zhang, G. (2019). Rapid determination of soil classes in soil profiles using vis–NIR spectroscopy and multiple objectives mixed support vector classification. *European Journal of Soil Science*, 70(1):42–53.

Chen, S., Xu, D., Li, S., Ji, W., Yang, M., Zhou, Y., Hu, B., Xu, H., and Shi, Z. (2020). Monitoring soil organic carbon in alpine soils using in situ vis-NIR spectroscopy and a multilayer perceptron. *Land Degradation & Development*, 31(8):1026–1038.

Cheng, Q. and Wu, X. (2008). Multisource remote sensing archaeology of the Ancient Canal in Zhejiang, China. In *Remote Sensing for Environmental Monitoring, GIS Applications, and Geology VIII*, volume 7110, page 71101T. International Society for Optics and Photonics.

Comer, D. C. and Harrower, M. J. (2013). *Mapping Archaeological Landscapes from Space*. SpringerBriefs in Archaeology. New York, NY [u.a.] Springer, (2013).

Conrad, C., Mertz, M., and Mester, R. (2013). Contour-Relaxed Superpixels. In *Energy Minimization Methods in Computer Vision and Pattern Recognition*, pages 280–293. Springer, Berlin, Heidelberg.

Cutler, A., Wiener, M., and Liaw, A. (2018). randomForest: Breiman and Cutler's Random Forests for Classification and Regression. https://CRAN.R-project.org/package=randomForest.

Dalponte, M., Bruzzone, L., Vescovo, L., and Gianelle, D. (2009). The role of spectral resolution and classifier complexity in the analysis of hyperspectral images of forest areas. *Remote Sensing of Environment*, 113(11):2345–2355.

Daniel, G. E. (1976). *A hundred and fifty years of archaeology*. Harvard University Press.

Daniel, K. W., Tripathi, N. K., and Honda, K. (2003). Artificial neural network analysis of laboratory and in situ spectra for the estimation of macronutrients in soils of Lop Buri (Thailand). *Soil Research*, 41(1):47–59.

Davies, S., Lamb, H., and Roberts, S. (2015). Micro-XRF Core Scanning in Palaeolimnology: Recent Developments. In Croudace, I. and Rothwell, R., editors, *Micro-XRF Studies of Sediment Cores: Applications of a non-destructive tool for the environmental sciences*, Developments in Paleoenvironmental Research, pages 189–226. Springer Netherlands, Dordrecht.

De Reu, J., De Smedt, P., Herremans, D., Van Meirvenne, M., Laloo, P., and De Clercq, W. (2014). On introducing an image-based 3D reconstruction method in archaeological excavation practice. *Journal of Archaeological Science*, 41:251–262.

De Roo, B., Stal, C., Lonneville, B., De Wulf, A., Bourgeois, J., and De Maeyer, P. (2016). Spatiotemporal data as the foundation of an archaeological stratigraphy extraction and management system. *Journal of Cultural Heritage*, 19:522–530.

Debret, M., Sebag, D., Desmet, M., Balsam, W., Copard, Y., Mourier, B., Susperrigui, A. S., Arnaud, F., Bentaleb, I., Chapron, E., Lallier-Vergès, E., and Winiarski, T. (2011). Spectrocolorimetric interpretation of sedimentary dynamics: The new "Q7/4 diagram". *Earth-Science Reviews*, 109(1):1–19.

Dinov, I. (2018). Black Box Machine-Learning Methods: Neural Networks and Support Vector Machines. In Dinov, I., editor, *Data Science and Predictive Analytics: Biomedical and Health Applications using R*, pages 383–422. Springer International Publishing.

Dravida, S., Woods, J., and Shen, W. (1984). A comparison of image filtering algorithms. In *ICASSP '84. IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 9, pages 259–262.

Duchon, J. (1977). Splines minimizing rotation-invariant semi-norms in Sobolev spaces. In *Constructive Theory of Functions of Several Variables*, Lecture Notes in Mathematics, pages 85–100. Springer, Berlin, Heidelberg.

Duzgoren-Aydin, N. S., Aydin, A., and Malpas, J. (2002). Re-assessment of chemical weathering indices: case study on pyroclastic rocks of Hong Kong. *Engineering Geology*, 63(1):99–119.

Einstein, A. (1955). Autobiographisches. In *P.A. Schlipp, Albert Einstein als Philosoph und Naturforscher*, pages 1–35. Stuttgart.

Evans, J. and Murphy, M. (2020). rfUtilities: Random Forests Model Selection and Performance Evaluation. https://github.com/jeffreyevans/rfUtilities.

Fairchild, M. (2013). *Color appearance models*. John Wiley & Sons.

Fattovich, R. (1977). Pre-aksumite civilization of Ethiopia: a provisional review. *Proceedings of the Seminar for Arabian Studies*, 7:73–78.

Fattovich, R. (2012). The northern Horn of Africa in the first millennium BCE: local traditions and external connections. *Rassegna di Studi Etiopici*, 4:1–60.

Feyerabend, P. (1975). *Against Method*. New Left Books.

Feyerabend, P. (1983). *Wider den Methodenzwang*. Suhrkamp.

French, C. A. I. (2003). *Geoarchaeology in Action: Studies in Soil Micromorphology and Landscape Evolution*. Psychology Press.

Funt, B. and Bastani, P. (2014). Irradiance-independent camera color calibration. *Color Research & Application*, 39(6):540–548.

Gama, J. (2016). *colorscience: Color Science Methods and Data.*

Gerlach, I. (2012). Yeha: An Ethio-Sabaean site in the highlands of Tigray (Ethiopia). *New research in archaeology and epigraphy of South Arabia and its neighbors*, pages 215–240.

Gerlach, I. (2013). Cultural Contacts between South Arabia and Tigray (Ethiopia) during the Early 1st Millennium BC. Results of the Ethiopian-German Cooperation Project in Yeha. *Zeitschrift für Orientforschung*, 6:254–277.

Gerlach, I. (2015). Migration Processes on the Northern Horn of Africa in the early first millenium BC. *Arabian and Islamic studies. A collection of papers in honour of Mikhail Borishovich Piotrovskij on the occasion of his 70th birthday*, pages 100–120.

Gerlach, I. (2017). Neue Forschungen zur äthio-sabäischen Kultur. *In kaiserlichem Auftrag: die Deutsche Aksum-Expedition 1906 unter Enno Littmann. Ethnographische, kirchenhistorische und archäologisch-historische Untersuchungen, Forschungen zur Archäologie außereuropäischer Kulturen*, (3):3.

Gerlach, I. (2018). Zum äthio-sabäischen Kunsthandwerk des frühen 1. Jahrtausends v. Chr. *Hauptsache Museum. Der Alte Orient im Fokus. Festschrift für Ralf-B. Wartke, marru 6*, pages 229–252.

Gill, D., Shomrony, A., and Fligelman, H. (1993). Numerical Zonation of Log Suites and Logfacies Recognition by Multivariate Clustering. *AAPG Bulletin*, 77(10):1781–1791.

Goetz, J. N., Brenning, A., Petschko, H., and Leopold, P. (2015). Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling. *Computers & Geosciences*, 81:1–11.

Goldberg, P., Holliday, V. T., and Ferring, C. R. (2013). *Earth sciences and archaeology*. Springer Science & Business Media.

Grimm, E. C. (1987). CONISS: a FORTRAN 77 program for stratigraphically constrained cluster analysis by the method of incremental sum of squares. *Computers & Geosciences*, 13(1):13–35.

Götze, A. (1912). Die vor- und frühgeschichtlichen Denkmäler des Kreises Westprignitz. *Kunstdenkmäler der Provinz Brandenburg*, 1(1).

Haburaj, V. (2019). UHD examineR. https://doi.org/10.5281/zenodo.3491697.

Haburaj, V., Krause, J., Pless, S., Waske, B., and Schütt, B. (2019). Evaluating the Potential of Semi-Automated Image Analysis for Delimiting Soil and Sediment Layers. *Journal of Field Archaeology*, 44(8):538–549.

Haburaj, V., Nykamp, M., May, J., Hoelzmann, P., and Schütt, B. (2020). On-site VIS-NIR spectral reflectance and colour measurements - a fast and inexpensive alternative for delineating sediment layers quantitatively? A case study from a monumental Bronze Age burial mound (Seddin, Germany). *Heritage*, 3(2):528–548.

Haburaj, V., Japp, S., Gerlach, I., Hoelzmann, P., Schütt, B. (2020). Coupling spectral imaging and laboratory analyses to digitally map sediment parameters and stratigraphic layers in Yeha, Ethiopia. *PLoS ONE*, 15(9): e0238894.

Haijun, Q., Xiu, J., Liu, Z., Maxime, D. I., and Shaowen, L. (2017). Predicting sandy soil moisture content with hyperspectral imaging. *International Journal of Agricultural and Biological Engineering*, 10(6):175–183.

Haralick, R. M., Shanmugam, K., and Dinstein, I. (1973). Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3(6):610–621.

Harris, E. C. (1979). *Principles of Archaeological Stratigraphy*. Academic Press, London.

Harris, E. C. (1989). *Principles of Archaeological Stratigraphy*. Academic Press, London, 2 edition.

Hartigan, J. A. and Wong, M. A. (1979). Algorithm AS 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28(1):100–108.

Hassan, F. A. (1978). Sediments in Archaeology: Methods and Implications for Palaeoenvironmental and Cultural Analysis. *Journal of Field Archaeology*, 5(2):197–213.

Heil, J., Marschner, B., and Stumpe, B. (2020). Digital photography as a tool for microscale mapping of soil organic carbon and iron oxides. *CATENA*, 193:104610.

Hengl, T., Heuvelink, G. B. M., Kempen, B., Leenaars, J. G. B., Walsh, M. G., Shepherd, K. D., Sila, A., MacMillan, R. A., Jesus, J. M. d., Tamene, L., and Tondoh, J. E. (2015). Mapping Soil Properties of Africa at 250 m Resolution: Random Forests Significantly Improve Current Predictions. *PLOS ONE*, 10(6):e0125814.

Hengl, T., Toomanian, N., Reuter, H. I., and Malakouti, M. (2007). Methods to interpolate soil categorical variables from profile observations: Lessons from Iran. *Geoderma*, 140(4):417–427.

Hijmans, R. J. (2017). raster: Geographic Data Analysis and Modeling. https://CRAN.R-project.org/package=raster.

Hijmans, R. J. (2019). raster: Geographic Data Analysis and Modeling. https://CRAN.R-project.org/package=raster.

Hill, J. and Schütt, B. (2000). Mapping Complex Patterns of Erosion and Stability in Dry Mediterranean Ecosystems. *Remote Sensing of Environment*, 74(3):557–569.

Hill, J., Udelhoven, T., Vohland, M., and Stevens, A. (2010). The Use of Laboratory Spectroscopy and Optical Remote Sensing for Estimating Soil Properties. In Oerke, E.-C., Gerhards, R., Menz, G., and Sikora, R. A., editors, *Precision Crop Protection - the Challenge and Use of Heterogeneity*, pages 67–85. Springer Netherlands, Dordrecht.

Hobley, E., Steffens, M., Bauke, S., and Kögel-Knabner, I. (2018). Hotspots of soil organic carbon storage revealed by laboratory hyperspectral imaging. *Scientific Reports*, 8(1):1–13.

Hoelzmann, P., Klein, T., Kutz, F., and Schütt, B. (2017). A new device to mount portable energy-dispersive X-ray fluorescence spectrometers (p-ED-XRF) for semi-continuous analyses of split (sediment) cores and solid samples. *Geoscientific Instrumentation, Methods and Data Systems*, 6(1):93.

Hornik, K. (2005). A CLUE for CLUster Ensembles. *Journal of Statistical Software*, 14(12). https://CRAN.R-project.org/package=clue.

Hoshyar, A., Al-Jumaily, A., and Hoshyar, A. (2014). Comparing the Performance of Various Filters on Skin Cancer Images. *Procedia Computer Science*, 42:32–37.

Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24(6):417–441.

Hsieh, P.-F., Lee, L. C., and Chen, N.-Y. (2001). Effect of spatial resolution on classification errors of pure and mixed pixels in remote sensing. *IEEE Transactions on Geoscience and Remote Sensing*, 39(12):2657–2663.

Hurst, V. J. (1977). Visual estimation of iron in saprolite. *GSA Bulletin*, 88(2):174–176.

Inglis, R. H., French, C., Farr, L., Hunt, C. O., Jones, S. C., Reynolds, T., and Barker, G. (2018). Sediment micromorphology and site formation processes during the Middle to Later Stone Ages at the Haua Fteah Cave, Cyrenaica, Libya. *Geoarchaeology*, 33(3):328–348.

Islam, K., Ploschner, M., and Goldys, E. M. (2017). Multi-LED light source for hyperspectral imaging. *Optics Express*, 25(26):32659–32668.

Japp, S. (2019). Yeha, Äthiopien. Archäologische Untersuchungen auf dem Kirchenvorplatz von Yeha. Die Arbeiten der Jahre 2013 bis 2018. *DAI e-Forschungsberichte*, pages 14–18.

Japp, S., Gerlach, I., Hitgen, H., and Schnelle, M. (2011). Yeha and Hawelti: cultural contacts between Saba and DMT — New research by the German Archaeological Institute in Ethiopia. *Proceedings of the Seminar for Arabian Studies*, 41:145–160.

Jarmer, T., Lavée, H., and Hill, P. S. . J. (2009). Using reflectance spectroscopy and Landsat data to assess soil inorganic carbon in the Judean Desert (Israel).

Jarmer, T. and Schütt, B. (1998). Analysis of iron contents in carbonate bedrock by spectroradiometric detection based on experimentally designed substrates. In *1st EARSeL Workshop on Imaging Spectroscopy. Remote Sensing Laboratories, University of Zurich, Switzerland*, pages 375–382.

Joshi, N., Baumann, M., Ehammer, A., Fensholt, R., Grogan, K., Hostert, P., Jepsen, M., Kuemmerle, T., Meyfroidt, P., Mitchard, E., Reiche, J., Ryan, C., and Waske, B. (2016). A Review of the Application of Optical and Radar Remote Sensing Data Fusion to Land Use Mapping and Monitoring. *Remote Sensing*, 8(1):70.

Jung, A., Vohland, M., and Thiele-Bruhn, S. (2015). Use of A Portable Camera for Proximal Soil Sensing with Hyperspectral Image Data. *Remote Sensing*, 7(9):11434–11448.

Kemper, T. and Sommer, S. (2002). Estimate of heavy metal contamination in soils after a mining accident using reflectance spectroscopy. *Environmental Science & Technology*, 36(12):2742–2747.

Kenyon, K. M. (1952). Excavations at Jericho, 1952. *Palestine Exploration Quarterly*, 84(2):62–82.

Kiekebusch, A. (1928). Das Königsgrab von Seddin. In *Führer zur Urgeschichte*, number 1. Augsburg.

Kooistra, M. J. and Kooistra, L. I. (2003). Integrated research in archaeology using soil micromorphology and palynology. *CATENA*, 54(3):603–617.

Lagacherie, P., McBratney, A., and Voltz, M. (2006). *Digital soil mapping: an introductory perspective*. Elsevier.

Landa, E. R. and Fairchild, M. D. (2005). Charting Color from the Eye of the Beholder. *American scientist*, 93(5):436–443.

Landesamt für Bergbau, G. u. R. (1905). Geologische Karte 1:25.000 (GK25) des Landes Brandenburg, Blatt Baek (2837).

Lark, R. M. (2005). Spatial analysis of categorical soil variables with the wavelet transform. *European Journal of Soil Science*, 56(6):779–792.

Lasaponara, R. and Masini, N. (2007). Detection of archaeological crop marks by using satellite QuickBird multispectral imagery. *Journal of Archaeological Science*, 34(2):214–221.

Legendre, P. (1993). Spatial Autocorrelation: Trouble or New Paradigm? *Ecology*, 74(6):1659–1673.

Lehnert, L. W., Meyer, H., and Bendix, J. (2017). hsdar: Manage, analyse and simulate hyperspectral data in R. https://CRAN.R-project.org/package=hsdar.

Leutner, B. and Horning, N. (2017). RStoolbox: Tools for Remote Sensing Data Analysis. https://CRAN.R-project.org/package=RStoolbox.

Linderholm, J., Geladi, P., Gorretta, N., Bendoula, R., and Gobrecht, A. (2019). Near infrared and hyperspectral studies of archaeological stratigraphy and statistical considerations. *Geoarchaeology*, 34(3):311–321.

Liu, H., Huang, M., Cui, G., Luo, M., and Melgosa, M. (2013). Color-difference evaluation for digital images using a categorical judgment method. *J. Opt. Soc. Am. A*, 30(4):616–626.

Liu, M. Y., Tuzel, O., Ramalingam, S., and Chellappa, R. (2011). Entropy rate superpixel segmentation. In *CVPR 2011*, pages 2097–2104.

Lodhi, V., Chakravarty, D., and Mitra, P. (2018). Hyperspectral Imaging for Earth Observation: Platforms and Instruments. *Journal of the Indian Institute of Science*, 98(4):429–443.

Madejova, J. and Komadel, P. (2001). Baseline studies of the clay minerals society source clays: infrared methods. *Clays and clay minerals*, 49(5):410–432.

Marshall, M., Schlolaut, G., Nakagawa, T., Lamb, H., Brauer, A., Staff, R., Ramsey, C., Tarasov, P., Gotanda, K., Haraguchi, T., Yokoyama, Y., Yonenobu, H., and Tada, R. (2012). A novel approach to varve counting using microXRF and X-radiography in combination with thin-section microscopy, applied to the Late Glacial chronology from Lake Suigetsu, Japan. *Quaternary Geochronology*, 13:70–80.

Marwick, B. (2017). Computational Reproducibility in Archaeological Research: Basic Principles and a Case Study of Their Implementation. *Journal of Archaeological Method and Theory*, 24(2):424–450.

Marwick, B., d'Alpoim Guedes, J., Barton, C., Bates, L., Baxter, M., Bevan, A., Bollwerk, E., Bocinsky, R., Brughmans, T., Carter, A., and others (2017). Open science in archaeology. *SAA Archaeological Record*, 17(4):8–14.

May, J. (2003). Der Fundplatz und die Umgebung des Grabhügels von Seddin als ortsfestes Bodendenkmal. In *Jürgen Kunow (Hrsg.), Das »Königsgrab« von Seddin in der Prignitz. Kolloquium anlässlich des 100. Jahrestages seiner Freilegung am 12. Oktober 1999*, number 9 in Arbeitsberichte zur Bodendenkmalpflege in Brandenburg, pages 7–11.

May, J. (2018a). Fokussieren, Positionieren, Schritthalten. Aspekte von Raum und Zeit am 'Königsgrab' von Seddin in der Prignitz. In *Das Ganze ist mehr als die Summe seiner Teile. Festschrift für Jürgen Kunow*, number 27 in Materialien zur Bodendenkmalpflege im Rheinland.

May, J. (2018b). Neue Forschung am 'Königsgrab' von Seddin. In Hansen, S. and Schopper, F., editors, *Der Grabhügel von Seddin im norddeutschen und südskandinavischen Kontext*, number 33 in Arbeitsberichte zur Bodendenkmalpflege in Brandenburg.

May, J. and Hauptmann, T. (2012). Das» Königsgrab «von Seddin und sein engeres Umfeld im Spiegel neuer Feldforschungen. In *D. Bérenger, J. Bourgeois, M.Talon, St. Wirth (Hrsg.), Gräberlandschaften der Bronzezeit. Internationales Kolloquium zur Bronzezeit*, number 51 in Bodenaltertümer Westfalens.

McBratney, A. B., Mendonça Santos, M. L., and Minasny, B. (2003). On digital soil mapping. *Geoderma*, 117(1):3–52.

Meier, T. and Tillessen, P. (2011). Von Schlachten, Hoffnungen und Ängsten. Einführende Gedanken zur Interdisziplinarität in der Historischen Umweltforschung. *Thomas Meier/Petra Tillessen (Hrsg.), Über die Grenzen und zwischen den Disziplinen*, pages 15–41.

Meister, J., Krause, J., Müller-Neuhof, B., Portillo, M., Reimann, T., and Schütt, B. (2017). Desert agricultural systems at EBA Jawa (Jordan): Integrating archaeological and paleoenvironmental records. *Quaternary International*, 434:33–50.

Melville, M. D. and Atkinson, G. (1985). Soil colour: its measurement and its designation in models of uniform colour space. *Journal of Soil Science*, 36(4):495–512.

Mesirov, J. (2010). Accessible Reproducible Research. *Science*, 327(5964):415–416.

Meyer, F. (1992). Color image segmentation. In *1992 International Conference on Image Processing and its Applications*, pages 303–306.

Millard, K. and Richardson, M. (2015). On the Importance of Training Data Sample Selection in Random Forest Image Classification: A Case Study in Peatland Ecosystem Mapping. *Remote Sensing*, 7(7):8489–8515.

Miller, C. E. (2001). Chemical principles of near-infrared technology. In Williams, P. and Norris, K., editors, *Near-infrared technology in the agricultural and food industries*, pages 19–37. The American Association of Cereal Chemists Inc.

Millward, A. (2011). Urbanisation viewed through a geostatistical lens applied to remote-sensing data. *Area*, 43(1):53–66.

Minasny, B. and McBratney, A. (2016). Digital soil mapping: A brief history and some lessons. *Geoderma*, 264:301–311.

Mittelstraß, J. (2003). *Transdisziplinarität - wissenschaftliche Zukunft und institutionelle Wirklichkeit*. Universitätsverlag Konstanz, Konstanz.

Morellos, A., Pantazi, X.-E., Moshou, D., Alexandridis, T., Whetton, R., Tziotzios, G., Wiebensohn, J., Bill, R., and Mouazen, A. M. (2016). Machine learning based prediction of soil total nitrogen, organic carbon and moisture content by using VIS-NIR spectroscopy. *Biosystems Engineering*, 152:104–116.

Morgan, C. and Wright, H. (2018). Pencils and Pixels: Drawing and Digital Media in Archaeological Field Recording. *Journal of Field Archaeology*, 43(2):136–151.

Munsell, A. H. (1929). *Munsell Book of Color*. Munsell Color Company, Baltimore.

Munsell Color Company (1994). *Munsell Soil Color Charts, Revised Edition*. Munsell Color Company, New York.

Mutuo, P. K., Shepherd, K. D., Albrecht, A., and Cadisch, G. (2006). Prediction of carbon mineralization rates from different soil physical fractions using diffuse reflectance spectroscopy. *Soil Biology and Biochemistry*, 38(7):1658–1664.

Nebes, N. (2010). Die Inschriften aus dem Almaqah-Tempel in Addi Akaweh (Tigray). *Zeitschrift für Orient-Archäologie*, 3:214–237.

Ng, W., Minasny, B., Montazerolghaem, M., Padarian, J., Ferguson, R., Bailey, S., and McBratney, A. (2019). Convolutional neural network for simultaneous prediction of several soil properties using visible/near-infrared, mid-infrared, and their combined spectra. *Geoderma*, 352:251–267.

Owen, A. J. (1995). Uses of derivative spectroscopy. *Agilent Technologies*, 8. Application Note.

Parker, A. (1970). An Index of Weathering for Silicate Rocks. *Geological Magazine*, 107(6):501–504.

Pebesma, E. (2012). spacetime: Spatio-Temporal Data in R. *Journal of Statistical Software*, 51(7):1–30.

Peng, R. (2011). Reproducible Research in Computational Science. *Science*, 334(6060):1226–1227.

Peña, M. A., Cruz, P., and Roig, M. (2013). The effect of spectral and spatial degradation of hyperspectral imagery for the Sclerophyll tree species classification. *International Journal of Remote Sensing*, 34(20):7113–7130.

Phillipson, D. W. (2009). The first millennium BC in the highlands of Northern Ethiopia and South–Central Eritrea: A reassessment of cultural and political development. *African Archaeological Review*, 26(4):257–274.

Piglia, R. (1992). *La ciudad ausente*. Editorial Sudamericana, Buenos Aires.

Post, J. L. and Noble, P. N. (1993). The Near-Infrared Combination Band Frequencies of Dioctahedral Smectites, Micas, and Illites. *Clays and Clay Minerals*, 41(6):639–644.

Pribyl, D. W. (2010). A critical review of the conventional SOC to SOM conversion factor. *Geoderma*, 156(3):75–83.

Price, J. and Velbel, M. (2001). An Assessment of Weathering Indices and Their Potential Applications to Heterogeneous Weathering Profiles and Paleosols. *Eleventh Annual V. M. Goldschmidt Conference*, 1.

Pyddoke, E. (1961). *Stratification for the Archaeologist*. Phoenix House, London.

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

R Core Team (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rees, W. G. (1990). *Physical Principles of Remote Sensing*. Number 1 in Topics in Remote Sensing. Cambridge University Press, Cambridge.

Renfrew, C. and Bahn, P. (1996). *Archaeology: Theories, Methods*. Thames & Hudson, London.

Retzlaff, R. (2015). *On the Potential of Small UAS for Multispectral Remote Sensing in Large-scale Agricultural and Archaeological Applications*. Doctoral Thesis, Universität Trier, Trier.

Robbins, J. (2018). Small Unmanned Aircraft Systems (sUAS). In *Horticultural Reviews, I. Warrington (Ed.)*, pages 33–71. John Wiley & Sons, Ltd.

Ropohl, G. (2009). *Allgemeine Technologie: Eine Systemtheorie der Technik*. Universitätsverlag Karlsruhe, Karlsruhe, 3 edition.

Ropohl, G. (2012). *Allgemeine Systemtheorie: Einführung in transdisziplinäres Denken*. edition sigma, Berlin.

Roskams, S. (2001). *Excavation*. Cambridge University Press, Cambridge, 1 edition.

Rungpichayapichet, P., Nagle, M., Yuwanbun, P., Khuwijitjaru, P., Mahayothee, B., and Müller, J. (2017). Prediction mapping of physicochemical properties in mango by hyperspectral imaging. *Biosystems Engineering*, 159:109–120.

Ruxton, B. (1968). Measures of the Degree of Chemical Weathering of Rocks. *The Journal of Geology*, 76(5):518–527.

Sankey, J. B., Brown, D. J., Bernard, M. L., and Lawrence, R. L. (2008). Comparing local vs. global visible and near-infrared (VisNIR) diffuse reflectance spectroscopy (DRS) calibrations for the prediction of soil clay, organic C and inorganic C. *Geoderma*, 148(2):149–158.

Sapirstein, P. and Murray, S. (2017). Establishing Best Practices for Photogrammetric Recording During Archaeological Fieldwork. *Journal of Field Archaeology*, 42(4):337–350.

Savitzky, A. and Golay, M. J. E. (1964). Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry*, 36(8):1627–1639.

Schenk, T. and Goldmann, T. (2004). Die Seddiner Kultfeuerreihe. Geomagnetische Prospektion am Königsgrab in der Prignitz. In *Archäologie in Berlin und Brandenburg 2003*, pages 57–59. Konrad Theiss, Stuttgart.

Schmid, T., Koch, M., DiBlasi, M., and Hagos, M. (2008). Spatial and spectral analysis of soil surface properties for an archaeological area in Aksum, Ethiopia, applying high and medium resolution data. *CATENA*, 75(1):93–101.

Schmidt, M., Leipe, C., Becker, F., Goslar, T., Hoelzmann, P., Mingram, J., Müller, S., Tjallingii, R., Wagner, M., and Tarasov, P. E. (2019). A multi-proxy palaeolimnological record of the last 16,600 years from coastal Lake Kushu in northern Japan. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 514:613–626.

Schoeneberger, P., Wysocki, D., Benham, E., and Staff, S. S. (2012). *Field book for describing and sampling soils, version 3.0*. Natural Resources Conservation Service, National Soil Survey Center. Lincoln, NE.

Schreiner, S., Buddenbaum, H., Emmerling, C., and Steffens, M. (2015). VNIR/SWIR Laboratory Imaging Spectroscopy for Wall-to-Wall Mapping of Elemental Concentrations in Soil Cores. *Photogrammetrie-Fernerkundung-Geoinformation*, 2015(6):423–435.

Schubert, H., Caballero Calvo, A., Rauchecker, M., Rojas-Zamora, O., Brokamp, G., and Schütt, B. (2018). Assessment of Land Cover Changes in the Hinterland of Barranquilla (Colombia) Using Landsat Imagery and Logistic Regression. *Land*, 7(4):152.

Schwertmann, U. and Cornell, R. M. (2008). *Iron Oxides in the Laboratory: Preparation and Characterization*. John Wiley & Sons, Weinheim.

Schütt, B. (2004). Reconstruction of Holocene weathering conditions from chemical character of playa-lake-sediments - a case study from central Spain. *Lecture Notes in Earth Sciences*, 102:5–30.

Segal, M. (2004). Machine Learning Benchmarks and Random Forest Regression. *UCSF: Center for Bioinformatics and Molecular Biostatistics*.

Shaw, G. and Burke, H. (2003). Spectral Imaging for Remote Sensing. *Lincoln Laboratory Journal*, 14(1):3–28.

Shepherd, J. D. and Dymond, J. R. (2003). Correcting satellite imagery for the variance of reflectance and illumination with topography. *International Journal of Remote Sensing*, 24(17):3503–3514.

signal developers (2014). signal: Signal processing. http://r-forge.r-project.org/projects/signal/.

Simonson, R. W. (1993). Soil Color Standards and Terms for Field Use—History of Their Development. *Soil Color*, SSSA Special Publication(Soil Color):1–20.

Skrabal, P. M. (2009). *Spektroskopie - Eine methodenübergreifende Darstellung vom UV- bis zum NMR-Bereich*. UTB.

Sorenson, P. T., Quideau, S. A., and Rivard, B. (2018). High resolution measurement of soil organic carbon and total nitrogen with laboratory imaging spectroscopy. *Geoderma*, 315:170–177.

Soriano-Disla, J. M., Janik, L. J., Viscarra Rossel, R. A., Macdonald, L. M., and McLaughlin, M. J. (2014). The Performance of Visible, Near-, and Mid-Infrared Reflectance Spectroscopy for Prediction of Soil Physical, Chemical, and Biological Properties. *Applied Spectroscopy Reviews*, 49(2):139–186.

Spinner, H. F. (1974). *Pluralismus als Erkenntnismodell*. Suhrkamp, Frankfurt.

Sponagel, H., Grottenthaler, W., Hartmann, K.-J., Hartwich, R., Janetzko, P., Joisten, H., Kühn, D., Sabel, K.-J., and Traidl, R. (2005). *Bodenkundliche Kartieranleitung. KA5*. Hrsg. von der Bundesanstalt für Geowissenschaften und Rohstoffe in Zusammenarbeit mit den Staatlichen Geologischen Diensten, Hannover.

Sprafke, T. (2016). *Löss in Niederösterreich - Archiv quartärer Klima- und Landschaftsveränderungen*. Würzburg University Press, Würzburg.

Stefanski, J., Mack, B., and Waske, B. (2013). Optimization of Object-Based Image Analysis With Random Forests for Land Cover Mapping. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(6):2492–2504.

Steffens, M. and Buddenbaum, H. (2012). Mapping the distribution of chemical properties in soil profiles using laboratory imaging spectroscopy, SVM and PLS regression. *EARSeL eProceedings*, 11(1):25–32.

Steffens, M. and Buddenbaum, H. (2013). Laboratory imaging spectroscopy of a stagnic Luvisol profile — High resolution soil characterisation, classification and mapping of elemental concentrations. *Geoderma*, 195–196:122–132.

Steffens, M., Kohlpaintner, M., and Buddenbaum, H. (2014). Fine spatial resolution mapping of soil organic matter quality in a Histosol profile. *European Journal of Soil Science*, 65(6):827–839.

Steger, S., Bell, R., Petschko, H., and Glade, T. (2015). Evaluating the Effect of Modelling Methods and Landslide Inventories Used for Statistical Susceptibility Modelling. In *Lollino G. et al. (eds), Engineering Geology for Society and Territory*, volume 2, pages 201–204. Springer, Cham.

Stenberg, B. and Viscarra Rossel, R. A. (2010). Diffuse Reflectance Spectroscopy for High-Resolution Soil Sensing. In Viscarra Rossel, R. A., McBratney, A. B., and Minasny, B., editors, *Proximal Soil Sensing*, Progress in Soil Science, pages 29–47. Springer, Dordrecht.

Stenberg, B., Viscarra Rossel, R. A., Mouazen, A. M., and Wetterlind, J. (2010). Visible and Near Infrared Spectroscopy in Soil Science. In Sparks, D. L., editor, *Advances in Agronomy*, volume 107, pages 163–215. Academic Press.

Stolt, M. and Baker, J. (1994). Strategies for Studying Saprolite and Saprolite Genesis. In *Whole Regolith Pedology*, volume 34 of *SSSA Special Publications*, pages 1–19. John Wiley & Sons.

Stoner, E. and Baumgardner, M. (1981). Characteristic variations in reflectance of surface soils. *Soil Science Society of America Journal*, 45(6):1161–1165.

Strahler, A. N. (1975). *Physical Geography*. John Whiley & Sons, New York, 4 edition.

Strauss, T. and Maltitz, M. J. v. (2017). Generalising Ward's Method for Use with Manhattan Distances. *PLOS ONE*, 12(1):e0168288.

Stutz, D., Hermans, A., and Leibe, B. (2018). Superpixels: An Evaluation of the State-of-the-Art. *Computer Vision and Image Understanding*, 166:1–27.

Sánchez-Marañón, M., Delgado, G., Delgado, R., Pérez, M. M., and Melgosa, M. (1995). Spectroradiometric and visual color measurements of disturbed and undisturbed soil samples. *Soil Science*, 160(4):291–303.

Sánchez-Marañón, M., García, P., Huertas, R., Hernández-Andrés, J., and Melgosa, M. (2011). Influence of Natural Daylight on Soil Color Description: Assessment Using a Color-Appearance Model. *Soil Science Society of America Journal*, 75(3):984–993.

Taubenböck, H., Esch, T., Felbier, A., Wiesner, M., Roth, A., and Dech, S. (2012). Monitoring urbanization in mega cities from space. *Remote Sensing of Environment*, 117:162–176.

Udelhoven, T., Emmerling, C., and Jarmer, T. (2003). Quantitative analysis of soil chemical properties with diffuse reflectance spectrometry and partial least-square regression: A feasibility study. *Plant and Soil*, 251(2):319–329.

Udelhoven, T. and Schütt, B. (2000). Capability of feed-forward neural networks for a chemical evaluation of sediments with diffuse reflectance spectroscopy. *Chemometrics and Intelligent Laboratory Systems*, 51(1):9–22.

Ulery, A. L. and Graham, R. C. (1993). Forest Fire Effects on Soil Color and Texture. *Soil Science Society of America Journal*, 57(1):135–140.

Umweltbundesamt, editor (2010). *Die Böden Deutschlands. Sehen, Erkunden, Verstehen*. Umweltbundesamt, Dessau-Roßlau.

Vaiphasa, C. (2006). Consideration of smoothing techniques for hyperspectral remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 60(2):91–99.

van Beijma, S., Comber, A., and Lamb, A. (2014). Random forest classification of salt marsh vegetation habitats using quad-polarimetric airborne SAR, elevation and optical RS data. *Remote Sensing of Environment*, 149:118–129.

Vincent, M. L., Kuester, F., and Levy, T. E. (2014). Opendig: Contextualizing the past from the field to the web. *Mediterranean Archaeology and Archaeometry*, 14(4):109–116.

Viscarra Rossel, R. A. and Behrens, T. (2010). Using data mining to model and interpret soil diffuse reflectance spectra. *Geoderma*, 158(1):46–54.

Viscarra Rossel, R. A., Behrens, T., Ben-Dor, E., Brown, D. J., Demattê, J. A. M., Shepherd, K. D., Shi, Z., Stenberg, B., Stevens, A., Adamchuk, V., Aïchi, H., Barthès, B. G., Bartholomeus, H. M., Bayer, A. D., Bernoux, M., Böttcher, K., Brodský, L., Du, C. W., Chappell, A., Fouad, Y., Genot, V., Gomez, C., Grunwald, S., Gubler, A., Guerrero, C., Hedley, C. B., Knadel, M., Morrás, H. J. M., Nocita, M., Ramirez-Lopez, L., Roudier, P., Campos, E. M. R., Sanborn, P., Sellitto, V. M., Sudduth, K. A., Rawlins, B. G., Walter, C., Winowiecki, L. A., Hong, S. Y., and Ji, W. (2016). A global spectral library to characterize the world's soil. *Earth-Science Reviews*, 155:198–230.

Viscarra Rossel, R. A., Cattle, S. R., Ortega, A., and Fouad, Y. (2009). In situ measurements of soil colour, mineral composition and clay content by vis–NIR spectroscopy. *Geoderma*, 150(3):253–266.

Viscarra Rossel, R. A., Walvoort, D. J. J., McBratney, A. B., Janik, L. J., and Skjemstad, J. O. (2006). Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma*, 131(1):59–75.

Vodyanitskii, Y. N. and Kirillova, N. P. (2016). Application of the CIE-L*a*b* system to characterize soil color. *Eurasian Soil Science*, 49(11):1259–1268.

Vogel, S., Märker, M., Rellini, I., Hoelzmann, P., Wulf, S., Robinson, M., Steinhübel, L., Di Maio, G., Imperatore, C., Kastenmeier, P., and others (2016). From a stratigraphic sequence to a landscape evolution model: Late Pleistocene and Holocene volcanism, soil formation and land use in the shade of Mount Vesuvius (Italy). *Quaternary international*, 394:155–179.

Vohland, M., Bossung, C., and Fründ, H. (2009). A spectroscopic approach to assess trace- heavy metal contents in contaminated floodplain soils via spectrally active soil components. *Journal of Plant Nutrition and Soil Science*, 172(2):201–209.

Volkan Bilgili, A., van Es, H. M., Akbas, F., Durak, A., and Hively, W. D. (2010). Visible-near infrared reflectance spectroscopy for assessment of soil properties in a semi-arid area of Turkey. *Journal of Arid Environments*, 74(2):229–238.

von Hoff, K. E. A. (1822). *Geschichte der durch Überlieferung nachgewiesenen natürlichen Veränderungen der Erdoberfläche*, volume I-V. Gotha.

Warburton, D. (2003). *Archaeological Stratigraphy: A Near Eastern Approach*. Recherches et publications, Neuchâtel.

Waske, B. and Benediktsson, J. (2007). Fusion of Support Vector Machines for Classification of Multisensor Data. *IEEE Transactions on Geoscience and Remote Sensing*, 45(12):3858–3866.

Wheeler, M. (1954). *Archaeology from the Earth*, volume 356. Oxford University Press, Oxford.

Wheeler, M. (1960). *Moderne Archäologie: Methoden und Technik der Ausgrabung*. Rowohlt, Reinbek b. Hamburg.

Willett, E. H. (1880). On Flint Workings at Cissbury, Sussex. *Archaeologia*, 45(2):337–348.

Wills, S. A., Burras, C. L., and Sandor, J. A. (2007). Prediction of Soil Organic Carbon Content Using Field and Laboratory Measurements of Soil Color. *Soil Science Society of America Journal*, 71(2):380–388.

Winterbottom, S. J. and Dawson, T. (2005). Airborne multi-spectral prospection for buried archaeology in mobile sand dominated systems. *Archaeological Prospection*, 12(4):205–219.

Woodcock, C. E. and Strahler, A. H. (1987). The factor of scale in remote sensing. *Remote Sensing of Environment*, 21(3):311–332.

Wu, D. and Sun, D.-W. (2013). Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: A review — Part I: Fundamentals. *Innovative Food Science & Emerging Technologies*, 19:1–14.

Wüstemann, H. (1974). Zur Sozialstruktur im Seddiner Kulturgebiet. *Zeitschrift für Archäologie*, 8:67–107.

Yao, J., Boben, M., Fidler, S., and Urtasun, R. (2015). Real-Time Coarse-to-Fine Topologically Preserving Segmentation. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2947–2955.

Zeeden, C., Krauß, L., Kels, H., and Lehmkuhl, F. (2017). Digital image analysis of outcropping sediments: Comparison to photospectrometric data from Quaternary loess deposits at Şanoviţa (Romania) and Achenheim (France). *Quaternary International*, 429:100–107.

Zhang, Y. and Hartemink, A. E. (2019a). Digital mapping of a soil profile. *European Journal of Soil Science*, 70(1):27–41.

Zhang, Y. and Hartemink, A. E. (2019b). A method for automated soil horizon delineation using digital images. *Geoderma*, 343:97–115.

Zhang, Y. and Hartemink, A. E. (2019c). Soil horizon delineation using vis-NIR and pXRF data. *Catena*, 180:298–308.

Zvoleff, A. (2016). glcm: Calculate Textures from Grey-Level Co-Occurrence Matrices (GLCMs). https://CRAN.R-project.org/package=glcm.

## Supporting information for chapter 4.2

- **Online Resource 4.1** comprises the R script, used to carry out the unsupervised classification of RGB image data. Data available here: `https://www.tandfonline.com/doi/suppl/10.1080/00934690.2019.1656321/suppl_file/yjfa_a_1656321_sm3187.r`.

- **Online Resource 4.2** comprises the R script, used to carry out the unsupervised classification of hyperspectral image data. Data available here: `https://www.tandfonline.com/doi/suppl/10.1080/00934690.2019.1656321/suppl_file/yjfa_a_1656321_sm3183.r`.

- **Online Resource 4.3** includes the RGB and hyperspectral image data used throughout chapter 4.2. Data available here: `https://www.tandfonline.com/doi/suppl/10.1080/00934690.2019.1656321/suppl_file/yjfa_a_1656321_sm3182.zip`

- **Online Resource 4.4** contains a detailed profile description of the two examined profiles. Data available here: `https://www.tandfonline.com/doi/suppl/10.1080/00934690.2019.1656321/suppl_file/yjfa_a_1656321_sm3179.pdf`

- **Online Resource 4.5** contains detailed information on the processing workflow applied in chapter 4.2. Data available here: `https://www.tandfonline.com/doi/suppl/10.1080/00934690.2019.1656321/suppl_file/yjfa_a_1656321_sm3175.pdf`

## Supporting information for chapter 5.2

- **Online Resource 5.1** includes the image data used throughout this study along with the R-scripts that were used for image analyses.

- **Online Resource 5.2** covers the elemental concentrations of the sediment samples retrieved from a p-ED-XRF, as well as documentation of the processing workflow of these data.

- **Online Resource 5.3** comprises the VIS-NIR spectra of all sediment samples, recorded with a spectroradiometer (350 - 2500 nm). Additionally, the documentation of the pre-processing workflow is included.

Online Resources 5.1 - 5.3 are available here: `https://www.mdpi.com/2571-9408/3/2/31/s1`.

# Supporting information for chapter 6.2

- **Online Resource 6.1** includes the detailed sedimentological data along with plots showing measurement results from ICP-OES and p-ED-XRF as well as XRD plots. Files available at: `https://doi.org/10.5281/zenodo.3906210`.

- **Online Resource 6.2** comprises R-Scripts for the regression analyses performed throughout the presented study along with the data and the produced plots. Files available at: `https://doi.org/10.5281/zenodo.3906216`.
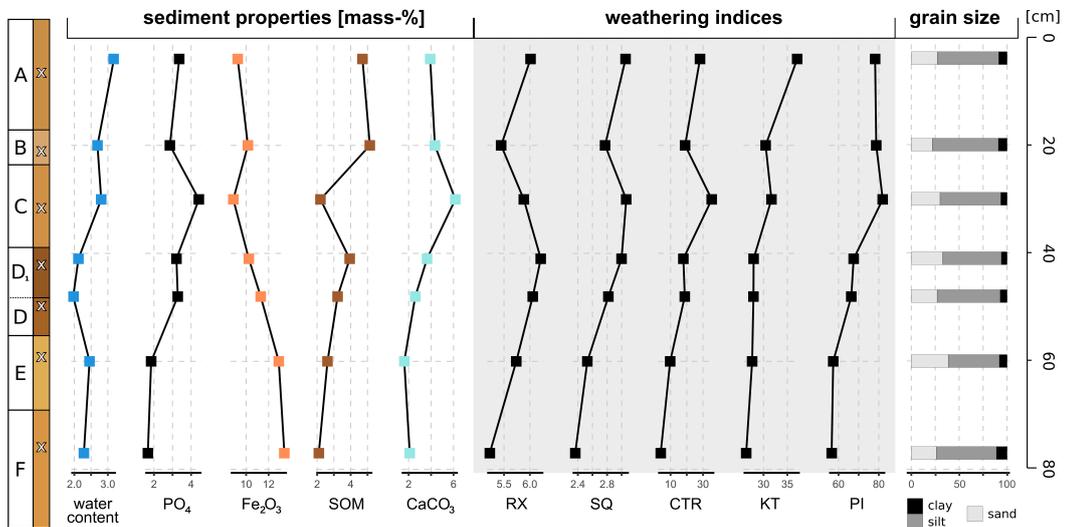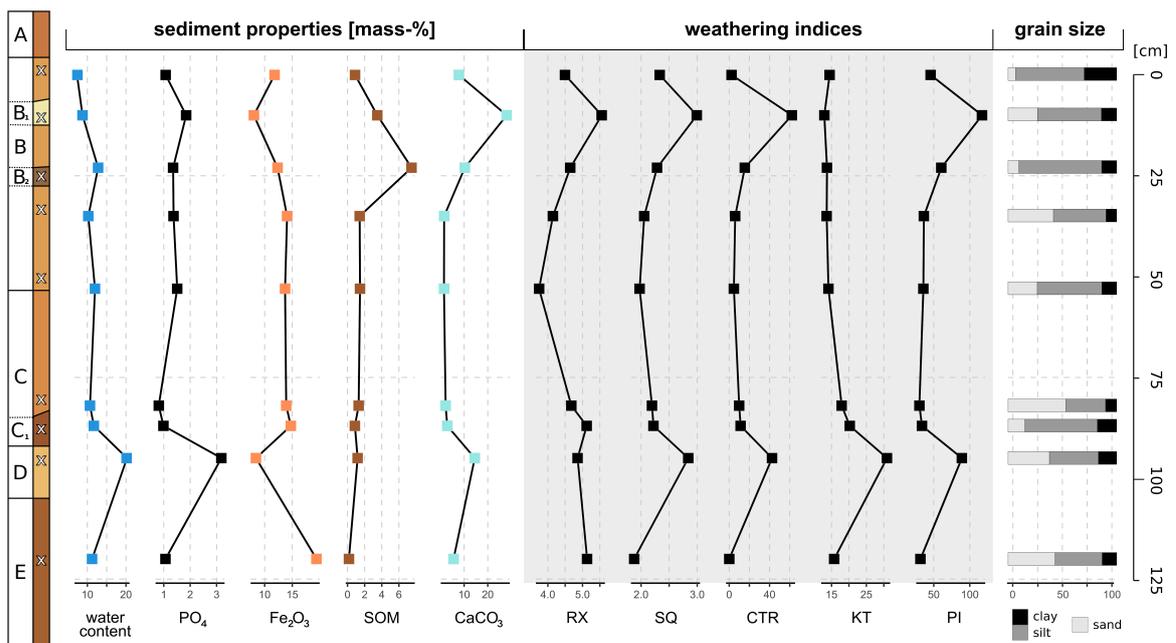
**Figure A.1:** Sedimentological record of profiles EI (top) and EVI (bottom). Water content of the sampled material was determined gravimetrically. Soil organic matter was calculated from quantitative measurements of total carbon (LECO TruspecCHN) and total inorganic carbon (Woesthoff Carmhograph C-16). Particle size distribution was measured using a laser diffraction particle size analyser (Beckmann-Coulter LS13 320). Element contents were measured using a portable energy-dispersive X-ray fluorescence spectrometer (Thermo Scientific Niton XL3t) and a inductively coupled plasma optical emission spectrometer (PerkinElmer Optima 2100 DV). Mineral composition was measured using an X-ray powder diffractometer (Rigaku MiniFlex 600). Weathering indices Ca-Ti-ratio (CTR), Ruxton-ratio (RX), Sesquioxide-ratio (SQ), K-Ti-ratio (KT) and Parker's Index (PI) were calculated. The methodology is described in more detail in the *Materials and methods* section.

# Curriculum vitae

For reasons of data protection, the curriculum vitae is not included in the online version.

# Eidesstattliche Erklärung

Hiermit erkläre ich, dass ich die Dissertation *Exploring Spectral Imaging as a Tool for Stratigraphic Analysis* selbstständig angefertigt und keine anderen als die von mir angegebenen Quellen und Hilfsmittel verwendet habe.

Ich erkläre weiterhin, dass die Dissertation bisher nicht in dieser oder in anderer Form in einem anderen Prüfungsverfahren vorgelegen hat.

Johann Joseph Vincent Haburaj

Berlin, den ............................