Hyperspectral imaging spectroscopy: a promising method for the biogeochemical analysis of lake sediments

Christoph Butz
Martin Grosjean
Daniela Fischer
Stefan Wunderle
Wojciech Tylmann
Bert Rein
Hyperspectral imaging spectroscopy: a promising method for the biogeochemical analysis of lake sediments

Christoph Butz, Martin Grosjean, Daniela Fischer, Stefan Wunderle, Wojciech Tylmann, and Bert Rein

University of Bern, Institute of Geography and Oeschger Centre for Climate Change Research, Erlachstrasse 9a, Bern CH-3012, Switzerland

University of Gdansk, Institute of Geography, Bazynskiego 4, Gdansk 80-952, Poland

GeoConsult Rein, Gartenstrasse 26-28, Oppenheim D-55276, Germany

Abstract. We investigate the potential of hyperspectral imaging spectrometry for the analysis of fresh sediment cores. A sediment-core-scanning system equipped with a camera working in the visual to near-infrared range (400 to 1000 nm) is described and a general methodology for processing and calibrating spectral data from sediments is proposed. We present an application from organic sediments of Lake Jaczno, a freshwater lake with biochemical varves in northern Poland. The sedimentary pigment bacteriopheophytin $a$ (BPhe $a$) is diagnostic for anoxia in lakes and, therefore, an important ecological indicator. Calibration of the spectral data (BPhe $a$ absorption $\sim$800 to 900 nm) to absolute BPhe $a$ concentrations, as measured by high-performance-liquid-chromatography, reveals that sedimentary BPhe $a$ concentrations can be estimated from spectral data with a model uncertainty of $\sim$10%. Based on this calibration model, we use the hyperspectral data from the sediment core to produce high-resolution intensity maps and time series of relative BPhe $a$ concentrations ($\sim$10 to 20 data points per year, pixel resolution $70 \times 70 \ \mu\text{m}^2$). We conclude that hyperspectral imaging is a very cost- and time-efficient method for the analysis of lake sediments and provides insight into the spatiotemporal structures of biogeochemical species at a degree of detail that is not possible with wet chemical analyses. © 2015 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.JRS.9.096031]

Keywords: hyperspectral imaging; image acquisition/recording; reflectance; nondestructive testing; sedimentary pigments; bacteriopheophytin.

1 Introduction

In situ imaging spectroscopy has great potential for fast, nondestructive, inexpensive, and high-resolution analysis of material compositions, supplementing established physical or chemical analytical methods. The identification of materials and substances according to their diagnostic spectral properties in the visible and infrared range has been extensively used at both remote and in situ scales. Typically, spectroscopic data from laboratory- or hand-held scanners are used for ground measurements to validate aerial or satellite remote sensing data or, for example, in the mining industry for geologic core logging. However, in contrast to airborne or satellite platforms, in situ scanners still mostly use point measurements instead of imaging methods, which limits the information about spatial variability in a given sample.

Recent advances in imaging spectrometry have opened a new field of scanning techniques for in situ or laboratory scales with hyperspectral resolution: hundreds of contiguous spectral bands at high spatial resolution (micrometer to nanometer scales) can be measured. This offers great opportunities in environmental research.

In environmental and earth-science research, biological, geochemical and physical properties of lake and marine sediments are widely used as records of past and recent environmental
changes. However, sample preparation and measurement with physical and chemical methods is generally very time consuming and expensive, which often limits the number of samples that can be analyzed. Therefore, nondestructive scanning and imaging methods have become increasingly important. In contrast to other well-established scanning methods like micro-x-ray-fluorescence (\(\mu\)XRF) or computer tomography, spectrometric methods are able to detect, based on their unique spectral fingerprints, chemical compounds and substances rather than chemical elements or physical properties (e.g., sediment density).

Spectroscopic methods using reflectance in the “visual to near-infrared range” (VNIR) have been used for decades for color logging (Munsell, CIELAB or other XYZ-based color spaces) on geologic sediment samples. The first applications for specific material identification and interpretation for different species were developed by Refs. References measured spectral data directly from fresh marine or lake sediment cores using point measurement photospectrometers (GretagMcBeth Spectrolino or ASD FieldSpec Pro). They derived spectral indices from reflectance spectra, which they could quantitatively calibrate to concentrations of chlorophyll \(a\) (Chl \(a\)), chlorins, lutein, and organic carbon. Furthermore, relative clay contents (illite, chlorite, and glauconite) could be estimated. All of these substances provide valuable information about environmental changes.

Following this approach, several authors successfully applied this method on recent freshwater lake sediments and demonstrated that, in specific cases, spectral data from point-measurements (2-mm sensor fields) along a sediment core could be directly calibrated to meteorological time series. In these cases, spectral data could be used as high-resolution proxies for quantitative climate reconstructions. In this study, we use a VNIR imaging spectrometer designed for use as a sediment core-scanning system and present a case study for spectral analysis of fresh lake sediments. First, we describe the technical equipment and the general methodology (workflow) that was used to acquire and process the spectral data. For illustration, the second part of this article presents a case study on organic sediments from a freshwater lake in Poland. We discuss the sample preparation for scanning and present a calibration procedure to calibrate spectral data to concentrations of specific substances, which were retrieved by established chemical analytical methods (e.g., high-performance-liquid-chromatography, HPLC). In our example, we show that concentrations of sedimentary bacteriopheophytin \(a\) (BPhe \(a\)) can be inferred at very high spatial resolution (70 \(\times\) 70 \(\mu\)m\(^2\)) for spatial maps and time series of BPhe \(a\) concentrations in a sediment core. BPhe \(a\) is mainly produced by bacteria under anoxic conditions and is thus most appropriate for the documentation of environmental change in freshwater environments.

2 Hyperspectral Imaging of Lake Sediments

2.1 General Methodology (Workflow)

The general methodology for spectral analysis of lake sediments consists of four steps (see Fig. 1; for more details, refer to Secs. 2.3–2.5 and 3 case study):

1. Data acquisition [Sec. 2.3, Fig. 1(a)]: the camera settings are optimized (field of view (FOV), focus, frame rate, exposure time, and moving speed of sample tray). The sediment core (sample) surface is cleaned and prepared. A scan of the white and dark standards and the sample is performed.

2. Data preprocessing [Sec. 2.4, Fig. 1(b)]: spectral data of the sediment core are calibrated to a white and dark standard. Regions of interest (ROIs) are defined and extracted from the calibrated image file of the core. A set of images [RGB, near infrared (NIR), color infrared (CIR) band combinations] is made for visual inspection of the data. Depending on the type of sample and quality of the scan, additional denoising can be performed on the dataset.

3. Data postprocessing [Sec. 2.5, Fig. 1(c)]: here, we use the Spectral Hourglass Wizard of the ENVI software for the extraction of the purest spectra in the sample (endmembers). After determination of the endmembers, spectral features within the endmembers are compared to diagnostic spectral indices known from the literature. These indices may then be modified and calculated for all spectra of the sample scanned. Unknown spectral
Fig. 1 General methodology for the hyperspectral analysis of lake sediments with the SCS. For details see Secs. 2 and 3. (a) Data acquisition, (b) pre-processing, (c) post-processing, and (d) output.
features may be marked and calculated for further investigation. Spectral endmembers can be compared to spectral libraries and used as classes for whole-sample classification.

(4) Output [Sec. 3, Fig. 1(d)]: typical types of data output include gray scaled and color mapped single-band images of spectral indices for visual inspection and as arrays, time series of spectral information along selected transects of the sediment core and statistics. Endmembers are stored as spectral libraries and graphs along with additional metadata and processing steps. Additional data analyses and outputs are made using R statistics.

2.2 Technical Instrument Description

Hyperspectral analyses are conducted using a sediment-core-scanning system (SCS, developed by Specim Ltd.) in combination with a Specim PFD-xx-V10E camera and a Schneider Kreuznach Xenoplan 1.9/35 lens. The SCS [Fig. 2(a)] consists of the spectral camera mounted above an illumination unit, a movable sample tray and a control unit connected to a personal computer. Images are measured in the VNIR spectrum from 400 to 1000 nm. The slit is 30 μm wide and the spectral resolution is 2.8 nm sampled at an interval of 0.78 nm or binned into 1.6, 3.2 or 6.4-nm intervals [Figs. 2(b) and 2(c)].33 Due to a small horizontal angle of view of the objective lens (~15.7 deg), errors related to lens aberration are insignificant. This permits the analysis of samples between ~45 and 120 mm widths and an on-sample spatial pixel resolution between 40 and 90 μm. Illumination is provided by a dome-like unit equipped with two arrays of 13 halogen lamps each, aligned across-track to evenly illuminate the sample. Illumination is indirect. The light is projected on a concave difusor plate and directed in such a way that it hits the sample at multiple angles.

The system uses a push broom technique. During a measurement, the tray with the sediment sample moves underneath the camera and illumination unit. The camera measures the light reflected from the sample line-by-line through a slit in the illumination unit. A PC interface (ChemaDAQ software, Specim Ltd.) controls the scanning system. Raw data are stored digitally on the hard drive. The raw data are compatible with ENVI’s binary format with separate header.
2.3 Data Acquisition

2.3.1 Sediment sample preparation

Sediment cores are typically retrieved in core liners of 1 to 1.5 m in length and 6 to 9 cm in diameter. In the lab, the cores are split lengthwise into two half cores and sometimes also subsampled with U-channels. Our system allows for sediment cores up to $12 \times 145 \text{ cm}^2$. In addition to fresh sediment cores, flat surfaces of epoxy-resin embedded and polished sediment slabs may be scanned.

After opening, the fresh sediment surface is cleaned with a sharp metal blade or a knife. The sediment surface is prepared in a way that means it is as flat as possible (allows for better focus) and the finest sediment structures are visible.

Specular reflection may be a problem for samples with high water content, especially with excess water on the surface. To minimize such effects of water reflection, wet sediments should be left to dry until surface water is sufficiently reduced. This process may take up to 24 h. This is particularly important for very dark sediments with generally low reflection.

2.3.2 Scanner and camera adjustments

Optimal settings for scanning depend on the sample type and its dimensions, as well as several hardware adjustments (e.g., camera focus), parameters and choices to be made by the operator.

Three steps are needed to focus the camera. First, a vertically moveable table [focus table, Fig. 2(d)] at the front of the sample tray is adjusted to the same height as the sample surface. The illumination unit is then positioned just above the focus table and the sample surface. Second, the FOV is set by changing the vertical position of the camera above the focus table. Third, tuning of the camera focus is done by manual adjustment of the lens. The focus table contains a grid of an alternating black and white pattern, which is used in combination with the computer monitor to focus the image. The focus table, therefore, corresponds to the focal plane of the later image.

Software-based settings (ChemaDAQ) include binning, frame rate, movement speed of the sample tray, exposure time, and scanning range. Spectral binning controls the spectral sampling rate. Typically, scanning is performed at the highest spectral sampling resolution ($\sim 0.78 \text{ nm}$) and subsequently reduced according to the scientific question. The standard frame rate for our equipment is 50 Hz due to the specifications of the PC interface. Thus, the maximum exposure time is 20 ms. According to the overall brightness of the sample, an appropriate exposure time is chosen. For dark samples, a higher exposure time is necessary to optimize the signal to noise ratio (SNR). The FOV determines the pixel size (FOV divided by the 1312 spatial pixels on the sensor in $x$-direction). Thus, in order to obtain a square pixel aspect ratio (1:1), both the speed of the sample tray and the exposure time have to be adjusted according to the FOV. To minimize potential influences from external light sources, laboratory windows are darkened and all other light sources are switched off prior to a scan. Complementary metal-oxide-semiconductor sensors (CMOS) are known to produce more noise with rising temperature. Therefore, we let the camera equilibrate for 10 min before performing the scans. A typical scan (FOV: 90 mm, spectral sampling: 0.78 nm, exposure time: 16 ms, sample length: 1.45 m) takes $\sim 8 \text{ min}$ to complete and produces $\sim 30 \text{ GB}$ of raw data.

To estimate mean error and signal strength, 50 measurements of the white and dark references are taken consecutively over a timespan of half an hour. The mean SNR (average of 50 white reference measurements) and mean standard error are then determined for each spectral channel [Figs. 3(a) and 3(b)]. SNR levels and the standard errors remain almost constant over this period. Standard error over time is $< 1\%$. SNR is highest in the range of 700 to 900 nm (ratio $\sim 200:1$) and drops on either side. For bands lower than 470 nm, the SNR drops below a ratio of 50:1. A low SNR could result in a fixed pattern noise (vertical striping) in the affected bands. The rapid drop of SNR in shorter wavelengths is due to several factors. Most importantly, the light source is
limited in this spectral range [Fig. 3(c)]. The sensor response and material of the lens also contribute to a weaker signal. Spectral bands lower than 405 nm are, therefore, not further considered due to generally low SNR. However, the SNR enables good performance of the camera in the mid to high range of the spectral channels (470 to 1000 nm).

2.4 Preprocessing of Spectral Data

First, the raw data are normalized to a white standard. Then, ROIs for relevant parts of the image are selected and subset to separate files. Typically, one ROI covers the entire sediment (complete core without non-sediment materials) and a second ROI defines a longitudinal transect. All procedures are standardized using ENVI/IDL programming interfaces.

White standard calibration is performed line by line following Eq. (1). Previously, the dark standard is subtracted from both raw data and the white standard. Optionally, a sample can be calibrated to an alternative white standard rather than the one created during the scan. To adjust for different exposure times between an alternative white standard and the raw data, a correction term is applied. The resulting data cube is a 32 bit floating point array of reflectance values in the range between 0 and 1.
where \( dc_{\text{raw}} \) = raw data, \( df_{\text{av}} \) = dark reference averaged into a single frame/scan line, \( wf_{\text{av}} \) = white reference averaged into a single frame/scan line, and \( t_{\text{int}} \) = integration time/exposure time. Analysis of sediment sample spectra requires that all nonrelevant parts of the image (e.g., scale bars, header, and core-liner) are removed. Typically, two subsets are created: one subset including the entire sediment core [Fig. 4(a), green box] and one smaller subset along an undisturbed transect in the \( y \)-direction, which is used for spectral analysis [Fig. 4(a), red box]. The smaller subset serves two purposes. First, sediment surfaces are not homogenous; therefore, a small transect allows one to select an undisturbed section. Second, the calculation times are substantially reduced if performed on a smaller region. All spectral bands below 405 nm are cut off during the subsetting procedure because of the generally low SNR. For visual inspection of the data, several tiff images of true color (RGB, R: 640 nm, G: 545 nm, B: 460 nm), CIR (R: 860 nm, G: 650 nm, B: 555 nm), and NIR (R: 900 nm, G: 800 nm, B: 700 nm) band combinations with different standardized contrast stretches are created [e.g., Figs. 4(b)–4(d)]. These images are used to accentuate weak changes in the sediment which would be hard to spot otherwise.

The lower bands (405 nm to ~470 nm) still show a fixed pattern noise (vertical striping) from the camera after normalization. In order to improve the image quality, a linear destriping method may be applied on the images (Fig. 5). Here, we use a linear destriping algorithm based on standard deviations and variance between the spatial sensor pixels. Very bright and very dark areas (e.g., water reflection, specular reflection, cracks, or holes) are masked prior to calculating image statistics. Typically, destriping is only performed on images used for visual inspection (e.g., RGB tiffs) in order to achieve a homogenous optical appearance. However, in some cases, destriping of the data bands may also improve the quality of the spectral analysis.

### 2.5 Postprocessing of Spectral Data

#### 2.5.1 Spectral analysis

Spectral analysis is performed on the small cross section [Fig. 4(a), red box] using the “Spectral Hourglass Wizard” of the ENVI software package (Exelis Visual Information Solutions, Boulder, Colorado). This procedure first reduces the spectral dimensionality of the dataset by a minimum noise fraction transformation and then spatially by a pixel purity index. A subsequent cluster analysis is performed and the average spectra of the clusters are calculated. Preclustered outputs from the ENVI software are usually checked and adjusted manually. The resulting spectra depict the endmembers of the spectral dataset. These can be compared to spectral libraries for material identification and can be used to classify the entire dataset using, for example, a spectral angle mapper.

#### 2.5.2 Spectral indices

In addition to spectral classification, indices may be calculated for spectral features (e.g., absorption bands) abundant in the sediment (as depicted by the endmembers) based on the diagnostic spectral properties of known substances (e.g., Chl \( a \)). The spectral endmembers aid in the decision to select appropriate indices. Diagnostic absorption bands can be used to make inferences about the identity and relative concentration of a substance. We follow the method of Ref. 23 for the calculation of spectral indices for absorption bands. A weighted average is calculated for two bands located at both ends of the absorption feature and a ratio is calculated between the weighted average and the absorption band minimum [see Eq. (2)]. We use continuum removal on the spectral endmembers in order to select the most appropriate bands for the calculation. This results in an index of relative absorption band depths (RABD).

\[
\text{RABD Feature MIN} = \frac{X \cdot R_{\text{Left}} + Y \cdot R_{\text{Right}}}{X + Y} / R_{\text{Feature MIN}},
\]

(2)
Fig. 4 (a) Overview of the sediment core from Lake Jaczno. Green and red boxes mark the different subsets. (b) True color (R: 640 nm, G: 545 nm, B: 460 nm) linear stretch of image subset in the boundary of the green box. (c) NIR band combination (R: 900 nm, G: 800 nm, B: 700 nm) linear stretch. (d) CIR band combination (R: 860 nm, G: 650 nm, B: 555 nm) linear stretch.
where \( R_{\text{ABD,Feature MIN}} \) = relative absorption band depth at absorption feature minimum, \( R_{\text{Left}} \) = reflectance at the start of the absorption feature, \( R_{\text{Right}} \) = reflectance at the end of the absorption feature, \( R_{\text{Feature MIN}} \) = reflectance at the minimum of the absorption feature, \( X \) = number of spectral bands between \( R_{\text{Feature MIN}} \) and \( R_{\text{Right}} \), and \( Y \) = number of spectral bands between \( R_{\text{Feature MIN}} \) and \( R_{\text{Left}} \).

Finally, the spectral indices and the other spectral outputs are ready to be used for interpretation and comparison to other datasets.

3 Case Study from Lake Jaczno, Poland

3.1 Introduction to the Case Study: Methodology for Calibration

In a case study from Lake Jaczno (54°17′E, 22°53′N, 163 m a.s.l. northeast Poland),\(^3\) we present an application of hyperspectral imaging on lake sediments and the methodology outlined above. Lake Jaczno is a postglacial mesotrophic freshwater lake with biochemical varved (annually laminated) sediments. The varves are composed of a light early summer layer rich in calcite and a dark late-summer and winter layer composed of organic and inorganic matter. One annual sediment layer is \( \sim 1 \) to 3 mm thick\(^3\) (Figs. 4 and 5). In our case study, we show that (1) bacteriochlorophyll \( a \) (BChl \( a \)) and its degradation product BPhe \( a \) present in the sediments of Lake Jaczno can be measured using hyperspectral imaging, (2) hyperspectral data can be calibrated to absolute BPhe \( a \) concentrations in sediments; and (3) high-resolution time series (\( \sim 20 \) to 30 data points per varve-year) and high resolution maps (70 \( \times \) 70 \( \mu \)m\(^2\)) of absolute pigment concentrations can be produced. Bacteriochlorophyll \( a \) and its diagenetic product BPhe \( a \) are photopigments produced by anoxic phototrophic bacteria (APB), indicating anoxia in the hypolimnion of a lake and a strong chemocline (chemically stratified water column). Bacteriopheophytin \( a \) is embedded and preserved in sediments under anoxic conditions and reveals information about the mixing regime and anoxia in lakes in the past.\(^4\) Such information is of great ecological interest. However, BPhe \( a \) is difficult to measure with wet chemical methods and the analysis is expensive.

Bacteriochlorophyll \( a \) and BPhe \( a \) show diagnostic absorption \textit{in situ} between \( \sim 800 \) and 900 nm.\(^4\)\(^4\)\(^2\) We hypothesize that the relative absorption around 845 nm might be related to BChl \( a \) and BPhe \( a \), rather than absorption due to mineral components, which may absorb light in this spectral range but do not possess as distinctive absorption bands. Bacteriochlorophylls and their derivative products are the only type of pigment known to absorb light in this range.
However, to identify the specific substance responsible for the absorption band (here, BPhe and RABD845) and to relate relative absorption depth to concentration values of the substance in the sediment matrix, calibration of the spectral data to a fully quantitative method is needed. Here, we use HPLC for the identification and quantification of BPhe. To do this, sediment subsamples were taken from the core and analyzed with wet chemical methods. Ideally, the subsamples are taken in a way that they are evenly distributed along the full range of final concentration values in the sediment core (stratified sampling), which is not known at this stage. However, the distribution of the relative absorption band depth of the investigated index (here, RABD845) can be used to make a selected choice for subsampling.

Hence, for the calibration of hyperspectral data to data from wet chemical or physical analyses, we propose a methodology with the following five steps (for details see Secs. 3.2–3.5):

1. A hyperspectral scan is performed (according to the methodology in Fig. 1) on the wet sediment, with subsequent normalization and spectral analysis. Spectral features, present in the endmembers, are compared to diagnostic spectral features of known substances of interest. User-defined spectral indices are calculated accordingly.

2. The map with the spectral index values is used to find suitable sites for subsampling the wet sediment. The subsampling sites are chosen to evenly cover the entire range of existing index values (stratified sampling). Suitable sites should be relatively homogeneous with regard to index values within this area. Subsampling of the wet sediment is then performed accordingly.

3. The subsampled sites of the sediment core are marked on the spectral index map (ROIs) and an average index value is calculated for each sediment sample. The subsamples, taken from the sediment core, are homogenized and hyperspectral scans are performed for the moist sediment subsamples and again for dry sediment subsamples after freeze drying. Averaged spectral indices are calculated for each subsample.

4. The analytes (substances of interest) are extracted from the sediment samples and analyzed quantitatively using established wet chemical or mineralogical methods (e.g., HPLC, gas chromatography, μXRF). The spectral index values are calibrated to the analytical results using regression models and validation techniques (e.g., bootstrapping, x-fold, and leave-one-out techniques).

3.2 Sediment Core Preparation

The 105-cm long and 9-cm diameter sediment core was split lengthwise into two half-cores. On opening, the sediment appeared to be almost black due to the metal reducing and anoxic conditions in which it was deposited. Also, the water content was very high. Therefore, the sediment surface was allowed to dry for ~24 h. Afterward, the oxidized surface developed a light brownish to reddish color, showing the annual laminations with pairs of bright and dark laminae [varves, Figs. 4(a)–4(d)]. Approximately 400 varves were counted in this core.

3.3 Acquisition of Spectral Data and Processing

Scanning was performed using a spectral sampling of 1.6 nm (binning of 2). The aperture of the lens was set to f/2.8. Thus, an exposure time of 18 ms was chosen based on the reflectance of the white reference to make use of the full radiometric range of the detector. Tray speed was adjusted to match the aspect ratio of 1:1. The raw data were then calibrated to the white and dark standards, and subsets for spectral analysis were taken [Fig. 4(a)]. A 2-mm wide transect in the center of the core was chosen for endmember analysis [Fig. 4(a), red box]. The position of this transect was selected based on visual evaluation of the sediment surface to avoid analysis of artifacts (e.g., holes in the sediment) that may influence the endmembers. Spectral endmembers were determined using the Spectral Hourglass Wizard [Fig. 6(a)]. The continua of the endmember spectra were then removed to find the boundaries of the absorption features [Fig. 6(b)].

Accordingly, the index RABD845 is calculated as follows:
The spatial distribution of RABD$_{845}$ index values across the sediment core was examined and 31 sites for sediment subsampling and chemical analysis (calibration) were determined according to the criteria described in Sec. 3.1, i.e., covering the entire range of index values present in the sediment core (Fig. 7).

3.4 Sediment Subsampling

Some sedimentary pigments tend to degrade soon after exposure to sunlight; samples were, therefore, only taken and processed in darkened rooms under indirect artificial light sources. For sampling, the sediment surface was carefully cleaned with a knife. The fresh material from the undisturbed interior of the sediment core was extracted, homogenized, filled into small, open sample boxes and scanned using the SCS (settings: FOV of 45 mm, fully opened aperture and 20-ms exposure time). Next, the samples were freeze-dried and scanned for a second time using the same settings.

$$\text{RABD}_{845} = \left( \frac{34 \times R_{790.26\,\text{nm}} + 34 \times R_{899.64\,\text{nm}}}{68} \right) / R_{844.73\,\text{nm}}. \tag{3}$$

The spatial distribution of RABD$_{845}$ index values across the sediment core was examined and 31 sites for sediment subsampling and chemical analysis (calibration) were determined according to the criteria described in Sec. 3.1, i.e., covering the entire range of index values present in the sediment core (Fig. 7).
settings. Subsequently, mean values of the RABD$_{845}$ index were calculated for each sample from both wet and dry scans.

### 3.5 Pigment Extraction and HPLC Analysis

For pigment extraction from the sediment samples, we followed the method of Refs. 40 and 45 adapted in Ref. 30. Samples were dissolved in pure acetone, filtered, and evaporated. Next, the extract was diluted in HPLC solvent and measured using an Agilent 1200 Infinity HPLC system. A series of standards of BChl a and its stable diagenetic product BPhe a were used for peak identification and quantification of the HPLC chromatograms.

### 4 Results

#### 4.1 Spectral Endmembers and RABD$_{845}$ Index

Analysis with the Spectral Hourglass Wizard revealed 47 spectral endmembers derived from the 2-mm wide transect [Fig. 4(a), red box, Fig. 6(a)]. The spectra show two major absorption features due to sedimentary pigments, with local absorption band minima at 672 and 845 nm. Most of the spectral endmembers were similar in shape but differed in their brightness. The calculation of the RABD$_{845}$ is shown in Fig. 6(b).

The map of RABD$_{845}$ index values and the corresponding spectral profile [mean values along the transect; red box in Fig. 4(a)] are shown in Fig. 8. The layers of the core denote the temporal behavior of sedimentation; therefore, the spectral profile can be considered as a time series. Both the map and the time series show areas and distinct layers in the sediment core that have high...
Fig. 8 Graphical output of the RABD$_{845}$ values calibrated to high-performance-liquid-chromatography (HPLC)-inferred BPhe $\alpha$ concentrations for the sediment core of Lake Jaczno showing (a) the RGB image, (b) the map of the spectral index distribution (RABD$_{845}$), and (c) the time series [average of 27 horizontal pixels (2 mm width)] within the boundary of the red lines. RABD$_{845}$ values were converted into BPhe $\alpha$ concentrations using the linear model [Fig. 9(a)]. $n$ is the number of pixels of the time series, i.e., the number rows of the RABD$_{845}$ map, respectively.
RABD$_{845}$ values (dark green to green color), whereas in some other areas, the RABD$_{845}$ values are around or even below a value of 1 (blue color), i.e., no absorption band is present.

### 4.2 Calibration of RABD$_{845}$ Values to BPhe a Concentrations

Four sediment samples were taken from sites in which no pigments associated to the index RABD$_{845}$ were expected and found (HPLC analysis). The mean RABD$_{845}$ value from these four samples (RABD$_{845} = 0.994$) was used as the baseline for the calibration (0.0 μg/g BPhe a). These four values were not used in the regression model. The linear regression between RABD$_{845}$ values (measured directly on the fresh sediment core) and BPhe a concentrations derived from HPLC analyses of the remaining 27 samples used for calibration (colored boxes in Fig. 7) is shown in Fig. 9. Statistical analysis showed that BPhe a concentrations of up to $\sim$28 μg/g are best described by a linear regression [Fig. 9(a)], while an exponential regression is the most appropriate for all BPhe a values, including very high values up to $\sim$46 μg/g. Excluding the RABD$_{845}$ data points below the baseline (RABD$_{845} < 0.994$, pigment concentration of 0.0 μg/g), the range of the linear model (approximately $< 28$ μg/g BPhe a) is valid for $\sim$97% of all data points measured (Fig. 8, black line). The correlation between the spectral data (RABD$_{845}$) and BPhe a (HPLC) is $R = 0.94$ ($p < 0.001$) with a coefficient of determination of $R^2 = 0.89$ [Fig. 9(a)]. The root mean square error of prediction (RMSEP) using 10-fold, leave-one-out ($k$-fold) and bootstrap methods averages $\sim 2.9$ μg/g BPhe a, which represents an uncertainty of about 10%.

The exponential calibration model [Fig. 9(b)] has a correlation coefficient of $R = 0.92$ and a corresponding coefficient of determination of $R^2 = 0.84$, which are similar to the values of the linear model. The model uncertainty for the ln-transformed concentrations is $\sim 10%$. Given that the exponential model only marginally increases the performance of the calibration in terms of data range and the results are slightly inferior to the linear model, we will use the latter during the remainder of the study.

### 4.3 Spatial Distribution of BPhe a (RABD$_{845}$ Values)

Figure 10 provides a detailed insight into the spatial distribution of RABD$_{845}$ values (BPhe a concentrations) within a very small 8-mm thick sediment subsection comprising three varve years. The pixel resolution of 70 μm allows for $\sim$30 spectra (vertically) per varve ($\sim$2-mm thick sediment layer) or per year. The RGB image [Fig. 10(a)] shows the sediment laminae couplets of the individual varves with the bright spring/early summer layers (mainly calcite) and

![Fig. 9](image-url) (a) Linear and (b) exponential regression model and calibration statistics between RABD$_{845}$ values and HPLC-inferred BPhe a concentrations from Lake Jaczno. The linear model is valid for concentrations up to $\sim$28 μg/g, the exponential model includes values up to $\sim$46 μg/g [$\sim 3.83 \ln(\mu g/g)$]. Dashed black lines are confidence intervals at 95% for the regression function and dashed green lines are confidence intervals at 95% for predicted values.
the dark organic rich late summer/winter layers. The spatial distribution map of the RABD$_{845}$
values [Fig. 10(b)] displays very high values (green color) in the lower part of the organic rich
dark layers (corresponding to summer and fall), low values (blue colors) in the upper part of the
dark organic rich sediment layers (corresponding to winter), and the light calcite rich spring/early
summer layer. The time series of RABD$_{845}$ values in Fig. 10(c) shows the mean values of all
points between the two red lines (∼2-mm wide) in Fig. 10(b). The BPhe $\alpha$ concentrations across
the three varve years show the same distribution pattern (generally, high values in the summer
and fall layer and low values in the winter layer and during calcite precipitation in spring/early
summer). They also show that the subvarve structure (i.e., the BChl $\alpha$ production and sedimen-
tation in individual years) is very different from year to year in amplitude and shape.

5 Discussion

5.1 SCS-VNIR Imaging and Methodology

Spectroscopic methods using reflectance in the VNIR range with direct in situ measurements on
the fresh sediment cores have been widely and successfully used for the biogeochemical analysis
of lake and marine sediments. These authors used point measurement photospectrometers typically with a spectral resolution of 10 nm, a spatial resolution (sensor field) of 2 to 8 mm and manual operation. The SCS, in comparison, improves the spatial resolution by about
two orders of magnitude (pixel size up to 40 $\mu$m) and has a better spectral resolution (2.8 nm).
The disadvantage is that the SNR is, in general, lower compared to the GretagMcBeth
Spectrolino, for example, especially in the shorter wavelength range (400 to 470 nm). This
limits the detection of ecologically important sedimentary substances (e.g., carotenoids, lutein,
etc.) that absorb in this shorter range. The SNR of the SCS is good between 480 and 900 nm,
which is ideal for detecting photopigments such as chlorophylls, bacteriochlorophylls, and their
diagenetic products (chlorins). Additionally, limited inferences may be made about the abundance
of certain clay minerals (e.g., chlorite, illite) in comparison to the presence of organic matter
and possibly other substances that still need to be explored and calibrated.

In comparison to established wet chemical analytical or mineralogical methods (HPLC, GC,
XRD, among others), hyperspectral VNIR imaging with the SCS is very fast, inexpensive
and allows for the analysis of structures and spatial distributions of substances at the submillim-
eter scale. The spatial resolution of data points generated by the hyperspectral SCS imaging

Fig. 10 Close-up of three varve years (38 to 46 mm sediment depth) with microstructures from
Lake Jaczno showing (a) the RGB image, (b) the RABD$_{845}$ index map, (c) and the time series of
the RABD$_{845}$ index values. The RABD$_{845}$ values are calibrated to BPhe $\alpha$ concentrations using the
linear model. The time series is the average of the index values of the rows between the two red
lines in (b). The light laminae in the RGB picture (a) represents the calcite layers (early summer sediments), the lower part of the dark sediment layers on top of the white layer represents the late summer and fall sediments, the upper part of the brown layer represents the winter sediments. Dashed black lines are a visual reference for sediment layers and RABD$_{845}$ values in the image and the graph.
“high resolution” with wet chemical techniques (e.g., for photopigment analysis by HPLC).\textsuperscript{30,40} The methodology outlined in Sec. 2 of this article is customized for scientists not familiar with remote sensing technologies, enabling them to acquire high-quality data and generate standard output products for spectral indices and sediment proxies that are well established in the environmental and paleoclimate literature (Refs.\textsuperscript{23, 25, 30}, and references therein).

5.2 Species Identification and Calibration of Spectral Indices

One of the major challenges that remains is the attribution of spectral properties (i.e., indices, endmembers) of lake sediments to specific substances present in the sediment matrix. Lake sediments can contain a large number of organic and inorganic substances,\textsuperscript{13} all of which influence the spectral reflectance (i.e., mixing of spectral signals). This means that diagnostic absorption bands only occur in rare cases and only for a few dominant substances (e.g., chlorophylls, bacteriochlorophylls, etc.). Thus, calibration of spectral properties with established quantitative and specific chemical and physical methods is essential. One problem is that VNIR spectroscopy takes a two-dimensional picture of the sediment surface, while chemical analyses are performed on a three-dimensional volumetric sample. This is particularly the case in sediments with very high spatial variability, such as varved sediments, as demonstrated in our case study. Consequently, accurate sampling of the sediments used for calibration is particularly important. Sections that are as homogenous as possible and have sufficient sample mass to perform the chemical analyses are ideal.

The calculation of the spectral indices is another important factor and the method should be carefully chosen. Pigment concentrations can be either derived from RABD or from relative absorption band areas (RABA).\textsuperscript{23,32} Either way, the selection of the bands may influence the result to a certain degree. In this study, we have chosen to use unfiltered reflectance data with a proven RABD algorithm as a basic approach. The bands used for calculations have been carefully selected using endmember analysis and continuum removal. However, further preprocessing of the input data (e.g., spectral smoothing) or the use of a RABA approach may enhance the calibration.

The performance of the linear calibration model for RABD\textsubscript{845} values and BPhe \textit{a} concentrations [Fig. 9(a)] is comparable to previous studies where different substances (i.e., \textit{chl} \textit{a}, chlorin, lutein, \textit{C}\textsubscript{org} and mineral group ratios) were calibrated to reflectance spectroscopy data from lake and marine sediments.\textsuperscript{21,23,32,49} Comparison of the regression models showed that the linear model (valid for BPhe \textit{a} concentrations < 28 \textmu g/g) performed slightly better than the exponential model (valid for BPhe \textit{a} concentrations < 46 \textmu g/g; Fig. 9).

We used spectral data that were measured directly on the fresh sediments. At least in the case study presented here, the calibration statistics did not improve if spectral data from homogenized wet or dry samples (step 3 of the calibration methodology; Sec. 3.1) were used instead of the spectral data from the sediment core image. Although this needs further testing with other lake sediments, it suggests that \textit{in situ} spectral measurements made directly on fresh sediments are suitable for calibration purposes.

5.3 Sedimentary BPhe \textit{a} Distribution in Lake Sediments (Our Case Study)

The calibration of RABD\textsubscript{845} to BPhe \textit{a} (HPLC) revealed that concentrations of BPhe \textit{a} in the Lake Jaczno sediment core could be estimated from hyperspectral data with an uncertainty of just 10\%. Figures 8 and 10 provide a detailed insight into the BPhe \textit{a} distribution along the entire sediment core (m-scale) and also within individual varves (submillimeter scale). Since one varve represents 1 year of sedimentation within the lake, a pixel resolution of 70 \textmu m allows analysis of seasonal patterns. The distribution of BPhe \textit{a} concentrations within one varve (Fig. 10) is consistent with the ecological conditions in the lake during the annual cycle. BPhe \textit{a} is the diagenetic product of Bchl \textit{a}, which is primarily produced by APB. APBs mainly develop in stratified lakes with seasonal sequences of hypolimnetic anoxia, which requires high primary production in the epilimnion (meso- to eutrophic lakes) and strong seasonal or semipermanent temperature and chemical stratification (di- or meromictic conditions with a strong thermo- and chemocline).\textsuperscript{42,50,51} Typically, lakes in northeast Poland mix twice a year in spring and fall when the
chemocline is gone and oxic conditions prevail in the entire water body. As a consequence, BPheα production, its sedimentation and concentration in the sediments are low in the spring layer and early summer layer. The early summer layer is clearly defined in the varves by large amounts of calcite, which contribute to the low BPheα concentration in the early summer sediments due to very fast precipitation (matrix effect). The situation is reversed in late summer when the lake is stratified, anoxia in the hypolimnion is well developed, production and sedimentation of BPheα are very high and thus, values in the organic-rich dark sediment layer above the light calcite layer are very high. In winter, photosynthesis is generally low under ice cover and BPheα concentrations in the winter layer (upper part of the organic-rich dark sediments) are low, despite the establishment of hypolimnetic anoxia.

Generally, the analysis of the three consecutive varves in Fig. 10 also shows that hyperspectral SCS data are able to depict differences in the structure and amplitude of BPheα concentrations from year to year. Short-term variability of stratification is often controlled by meteorological phenomena (mainly temperature and wind). However, the full potential of hyperspectral SCS data to interpret these subvarve structures in terms of climate reconstructions is yet to be investigated.

6 Conclusions and Outlook

This study has introduced a sediment core scanner equipped with hyperspectral VNIR imaging spectroscopy as a promising nondestructive method to study specific biochemical components of lake sediments. The imaging technique is inexpensive and fast compared to established methods, but the hyperspectral information and data need to be calibrated and verified.

The methodology proposed for sample preparation, data acquisition and processing, and generation of a standard set of output (Sec. 2.1) is widely applicable in the study of marine and lake sediments. All procedures are standardized using ENVI/IDL programming interfaces and output data can be further processed using R or other software. This standard procedure is designed and customized to enable nonspecialists in remote sensing to use this technology and data.

The SNR ratio is best between 470 and 900 nm. This is not ideal for the detection of sedimentary carotenoids, which absorb at shorter wavelengths (400 to 450 nm, e.g., lutein) but is very suitable for the detection of chloropigments and possibly other substances that are yet to be explored.

Calibration of hyperspectral data to data obtained from specific quantitative chemical or physical methods remains a challenge. Most critical is the design of the calibration experiment, the selection of the sediment sampling sites, and the analytical work. The proposed methodology for the calibration was shown to be successful in our case study, and the calibration statistics of direct spectral measurements on the fresh sediment surface ($R^2 = 0.89$, RMSEP = 10%) are comparable to previous studies. While it requires further testing, this methodology may be widely applicable to other studies.

Our case study shows that hyperspectral imaging of lake sediments may provide information about important environmental variables at high spatial and temporal resolution back in time (70 × 70 μm²/pixel). In this study, this means tens of data points within 1 year of sediment deposition, which is about three orders of magnitude better than what is possible with wet chemical methods.

We identify three potential future research areas: (1) hardware development, in particular, the extension of the spectral range to include shorter wavelengths at a higher SNR (380 to 470 nm); (2) further testing with case studies from other types of sediments and matrices, and calibrating more substances; and (3) exploring the potential of high-resolution hyperspectral data from varved lake sediments to reveal information about subseasonal phenomena. Climate reconstructions for the past hundreds to thousands of years based on time series of subseasonally resolved sediment archives would open new perspectives in climate research.

Acknowledgments

This research has been funded through grants from the Oeschger Centre and the Swiss contribution to the enlarged European Union CLIMPOL (PSRP-086/2010). Additional funding has been provided by the Swiss National Science Foundation (200020-134945/1). We would like to
thank the following people for their help and contribution to this research project: Drs. Benjamin Amann, Pascal Küpfer, Krystyna Saunders, and Tobias Schneider.

References


52. A. Bonk et al., “Modern limnology, sediment accumulation and varve formation processes in Lake Żabińskie, northeastern Poland: comprehensive process studies as a key to understand the sediment record,” *J. Limnol.* 73(AoP), 358–370 (2014).

Biographies for the authors are not available.