High-resolution grain size distribution of sediment core with hyperspectral imaging

Kevin Jacq, Charline Giguet-Covex, Pierre Sabatier, Yves Perrette, Bernard Fanget, Didier Coquin, Maxime Debret, Fabien Arnaud

High-resolution grain size distribution of sediment core with hyperspectral imaging

Kévin Jacq[1,2], Charline Giguet-Covex[1], Pierre Sabatier[1], Yves Perrette[1], Bernard Fanget[1], Didier Coquin[2], Maxime Debret[3], Fabien Arnaud[1]

1 Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, EDYTEM, 73000 Chambéry, France

2 Laboratoire d'Informatique, Systèmes, Traitement de l'Information et de la Connaissance (LISTIC), Université Savoie Mont-Blanc, 74944 Annecy Le Vieux Cedex, France

3 Laboratoire de Morphodynamique Continentale et Côtière, Université de Rouen, UMR CNRS 6143, 76821 Mont-Saint-Aignan, France, Université de Caen, UMR CNRS 6143, 14000 Caen, France

Corresponding Author:

Kévin Jacq

UMR CNRS 5204 Environnements, Dynamiques et Territoires de la Montagne (EDYTEM)

Université Savoie Mont Blanc, Campus scientifique,

73376 Le Bourget du Lac cedex, France

kevin.jacq@univ-smb.fr

Abstract

The study of sediment cores allows for the reconstruction of past climate and environment through physical-chemical analysis. Nevertheless, this interpretation suffers from many drawbacks that can be overcome with the newest technologies. Hyperspectral imaging is one of these and allows a fast, high resolution, and non-destructive analysis of sediment cores. In this study, we use visible and near-infrared hyperspectral imaging to predict particle size fractions and distribution (PSD) at a resolution of 200 µm on a previously well-studied sediment core taken from Lake Bourget (Western Alps, France). These predictions agree with previous studies on this core. Then, the PSD was used to estimate sedimentary deposit sources using the PSD unmixing algorithm AnalySize. It permitted estimation of the contribution of five sources (micrite, small and large bio-induced calcite crystals, diatom frustules, detrital particles), which had previously been characterized. The spatial dimension allowed for laminae to be discretized and counted, in agreement with the age-depth model previously established. We then evaluated the particle size and spectral signatures of each of these annual laminae, hence characterizing their physico-chemical composition. These high-resolution data also allowed for estimation of the accumulation rate (cm/year) of each of the main sources in the laminated unit and inferring the trophic status and the presence of instantaneous events of the lake.

Keywords: Grain size distribution, Hyperspectral Imaging, Chemometrics, Visible and Near-Infrared Spectroscopy, End-Members

1. Introduction

Particle size analysis is widely used in sediment studies. Several methods exist, such as gravimetric sedimentation, sieving or laser diffraction. The last one is the most used for fine lacustrine sediment. Such laboratory analyses of sediment textures are time and material consuming, labor-intensive, destructive and have low spatial resolution (0.5-1 cm sampling-steps are typically run) and thus temporal resolution. Spectroscopic methods show great possibilities to estimate particle size fractions (clay, silt, sand) without any sampling (Viscarra Rossel et al., 2016; de Santana et al., 2018; Nanni et al., 2018). The classically used coarser fractions, i.e., silt and sand, seem the most complicated to predict in several studies (de Santana et al., 2018; Nanni et al., 2018). The physical properties of the sediment can be estimated with the spectra baseline that is affected by the grain size in the visible and near-infrared ranges (Wetzel, 1983). It was also done in remote sensing with multispectral imaging (Casa et al., 2013; Demattê et al., 2015; Nouri et al., 2017). Multi- and hyper-spectral imaging are methods in the interface of spectroscopy and imaging with the acquisition of an image in which each pixel is characterized by a reflectance spectrum. Recently developed hyperspectral core loggers open new possibilities for the high-resolution prediction of organic matter (Van Exem et al., 2018; Jacq et al., 2019b) and pigments (Butz et al., 2017) and could thus be used also for the estimation of particle size.

The particle size distribution (PSD) is even more informative than the particle fractions. Some studies show the potential of visible and near spectroscopy to estimate the PSD (Hermansen et al., 2017). To go further, the PSD allows for differentiation and quantification of the different compounds of the sediment (deposition sources, processes), each of which is characterized by its PSD signature. Several algorithms have been proposed to estimate the relative contribution of these pure signatures in a sediment mixture (Heslop et al., 2007; Dietze et al., 2012; Paterson and Heslop, 2015; Yu et al., 2016;

Zhang et al., 2017). Van Hateren et al. (2017) argued that only a statistical comparative study of these estimated end-members and or a geological background could validate these signatures. The particle size end-member analysis of annually deposited sediments permits discrimination of allochthonous and autochthonous compounds to study varve formation and seasonal deposit variations over time (Żarczyński et al., 2019). The estimated end-members also allow for characterization of their relative abundance in each sample and therefore through time. It is also interesting to study instantaneous events, such as floods, to estimate their intensity for example.

This study aims to use hyperspectral imaging to estimate particle size fractions and distribution, then to estimate and characterize end-members. Partial least squares regression (PLSR) is one of the most used multivariate methods for the development of a quantitative predictive model with spectroscopic and physical-chemical data. PLSR is used to estimate particle size fractions and distribution in the same way as for organic matter (Jacq et al., 2019b). Two hyperspectral imaging sensors were used to cover the spectroscopic ranges from visible to near-infrared light (Vis-NIR: 400-2500 nm). A laser grain-sizer was used to estimate the particle size distribution and then after statistic treatment, the main fractions and other variables (D50, D90, skewness, kurtosis). The AnalySize Matlab function was used to unmix the PSD end-members and to estimate their abundance along the core. The proposed methodology is applied to a Lake Bourget sediment core and is compared with previous studies (Giguet-Covex et al., 2010; Jenny et al., 2013, 2014). Then, the end-members were used to characterize the annual laminations (varves), to count them, to estimate their accumulation rate per year and their Vis-NIR signatures.

2. Site description

Lake Bourget is a hard-water lake, 18 km long and 2.8 km wide, located at an altitude of 231.5 m at the northwestern edge of the French Alps (Fig. 1). The rivers Leysse and Sierroz flow into Lake

Bourget, which then flows into the Rhone River by the Saviere Channel under normal hydrological conditions. However, during major floods, water from the Rhone River flows into Lake Bourget with the reversal of the water-current of this channel. The sediment is mainly composed of river-borne silicate, between 10-40%, and carbonate, between 60-90%. The ratio between these two compounds varies depending on past climate conditions (Arnaud et al., 2005, 2012; Debret et al., 2010; Giguet-Covex et al., 2010; Jenny, 2013; Jenny et al., 2014).

The location of core LDB09-P101 (length = 54 cm, width = 9 cm) was selected in the northern basin of the lake (45°45'20.0"N 5°51'19.9"E, 145 m water depth). It was sampled for the IPER-RETRO program (ANR-08-VUL 005; Perga et al., 2015)). This site was selected because it contains both laminated sediments as well as interbedded deposits triggered by the Rhone River (Jenny, 2013).

3. Materials and methods

3.1. Sample preparation and measurement

The core was divided into two halves, one that was used for PSD and several destructive analyses, the other used for non-destructive analysis such as the hyperspectral imaging. This LDB09-P101 core is correlated to dated reference cores sampled in this lake, thanks to radionuclides (^{210}Pb, ^{137}Cs, ^{241}Am), historically known instantaneous events and varve counting (Jenny, 2013).

3.1.1.Particle size distribution

Particle size distribution (PSD) analyses were carried out on bulk sediments (5 mm thick, n = 18) using a Malvern Mastersizer 800S. This device uses the laser diffraction principle to estimate the particle diameter and estimate a distribution of them for 64 known classes between 0.06 µm and 800 µm. Then statistics allowed to calculate the mean particle size, several percentiles (e.g., median, D90, D99), mode, sorting, and particle size fraction (clay, silt, sand).

106 *3.1.2. Hyperspectral Imaging analyzes and registration*

107 A hyperspectral image (HSI) is an image with high spatial resolution, in which each pixel contains a
108 spectrum with a continuous spectral resolution.

109 The core was analyzed with two hyperspectral sensors (Specim Ltd., Finland) presented in Table 1
110 with the lens OLES22,5 at the M2C lab, University of Normandie-Rouen.

111 The sediment core must be as flat as possible for the analysis, so cleaning of the core was made, that
112 also revealed the finest sediment structures. Spectral calibration of the camera was made with a
113 spectralon reference to have reliable intensity. The spatial calibration was achieved by imaging a
114 known object to obtain its true shape and thus obtain squared pixels. The calibration was checked at
115 the end of the acquisition and no deviation was observed. More details about the acquisition
116 protocol can be found in Butz et al. (2015).

117 A composite sensor was created by down-sampling the VNIR sensor at 200 µm and a spatial
118 registration to fit spatially the two hyperspectral images. The registration was made with the
119 moisture content that can be recovered in the two sensors at 970 nm for the VNIR and at 1200 nm
120 for the SWIR (Ce Liu et al., 2011; Jacq et al., 2019a). The SWIR image was registered on the VNIR one,
121 because it has a lower resolution. Thus, a deformation model was estimated with these wavelengths
122 and used on all the SWIR spectral planes. Then the two registered hyperspectral images were put
123 together to create a composite sensor that covers the wavelengths between 400 nm and 2574 nm. It
124 was not possible to up-sample the SWIR sensor at 50 µm because the spatial details cannot be
125 estimated (Jacq et al., 2019a).

126 ## 3.2. Data Analysis

ENVI 4.8 was used to perform the data acquisition. Then, MATLAB (R2018a, MathWorks) was used for advanced processing and analyzing these data. PLSR was used to establish a relationship between the HSI and the PSD of the core to create predictive models and then to estimate the PSD in each pixel. Then AnalySize Matlab function was used to unmix the predicted PSD to find pure sources (Paterson and Heslop, 2015).

3.2.1.Predictive model estimation

A multivariate method was used to create quantitative predictive models with the particle size and the VNIR-SWIR spectra of the core. Several noisy bands due to the detector were removed on two sides of the spectra of each sensor; consequently 202 of 242 wavelengths were used. An autoscale spectral pre-processing was realized to reduce the matrix effect. Calibration spectra are median spectra of the particle size sampling areas.

PLSR was used to model the relation between the spectra and the reference particle size values. This method is based on the extraction of orthogonal predictor variables (also called latent variables, LV) corresponding to the maximum variability in the spectral bands used as predictors linked to one or several predicted variable(s) (Wold et al., 1984). Two main algorithms exist, called PLS1 and PLS2, the difference is the number of dependent variables to predict. PLS1 is used when a single variable is predicted. When there are several variables to predict, several PLS1 can be realized, or a single PLS2 if variables to predict are dependent. PLS1 was used to predict particle size fractions, whereas PLS2 to predict the PSD because the fine particle fractions are highly correlated and co-dependent. The spectra register some biological-physical-chemical properties of the sample. For a specific proxy, all these spectral information were not mandatories, so we decided to use a variable selection algorithm (random frog, genetic algorithm) with the PLSR. This allowed to find the most relevant wavelengths, decrease redundancy and to interpret the model chemically. More details about this

methodology can be found in Jacq et al. (2019b). Model accuracy was estimated with calibration and validation correlation coefficients and its uncertainty (Unc (%)). Another correlation is calculated between the median predictions of the sampling areas and the reference PSD to validate that the surface acquisition by hyperspectral methods contains the volume information of the particle size analyses ($r_{V/S}$).

3.2.2.Sedimentary deposition sources discrimination

Each PSD is a mixture of pure sub-PSD, that can be related to some specific sediment source. The Matlab function AnalySize (Paterson and Heslop, 2015) was used to unmixed each PSD and find these end-members. The assumption behind this function is a linear model equation (Eq. 1):

$$X = A \times S^T + E \quad \text{(Eq. 1)}$$

with X the reference PSD profile, A the estimated abundance, S^T the estimated PSD of the pure components and E the residual matrix.

The intensity of the PSD and its abundance should not be negative. Thus, non-negative spectra and concentrations constraints were used. Another constraint was added to have the sum of abundance and the intensity of S^T at 1.

AnalySize allows estimating end-members in two ways. One is called non-parametric end-member analysis (EMA) that estimates end-members empirically from the whole dataset. The other uses a parametric distribution (Weibull, skewed generalized Gaussian distribution) to fit the PSD. A parametric distribution was chosen to unmix the predicted PSD to overcome uncertainties of the model prediction that disturb the sum of Gaussian theoretical distributions.

4. Results

4.1. Lithology and particle size variations

171 The Lake Bourget sediment core presents three main units (Fig. 2b) that are mainly composed of clay
172 and silt fractions (Giguet-Covex et al., 2010).

173 Unit I, between 0 and 18 cm depth, presents millimeter-thick biochemical varves and corresponds to
174 perennial anoxic conditions at the lake-sediment interface, with a particle size mode near 22 µm
175 (Giguet-Covex et al., 2010; Jenny, 2013). This facies is related to the development of human-triggered
176 lake eutrophication. It is characterized by a triplet of laminae made of different components (Giguet-
177 Covex et al., 2010) (Supplementary materials 1, 2). The lowermost lamina (dark-brown after
178 oxidation) is deposited in spring and is mainly composed of diatom frustules and large bio-induced
179 calcite crystals. The intermediate lamina (whitish to ochre after oxidation) is deposited in summer. It
180 is composed of large and mostly small calcite crystals, micrite, few diatoms frustules and organic
181 matter. The uppermost grey lamina corresponds to the winter deposit, which mainly contains organic
182 matter and detrital particles. Some years a bloom of diatoms can occur in the fall. The presence of
183 the large bio-induced calcite crystals and a high amount of diatoms is characteristic of the
184 eutrophication period.

185 Unit II, between 18 cm and 26 cm depth, presents disturbed varves with a decreasing particle size
186 mode from 22 µm to 6 µm. Unit III, between 26 cm and 54 cm, is characterized by fuzzy laminated
187 structures and a particle size mode near 6 µm. In the two first units, there are also instantaneous
188 event deposits mostly triggered by floods (Jenny et al., 2014). One major instantaneous event occurs
189 between 13.5 cm and 14.5 cm (Fig. 2b).

190 Eighteen measurements of PSD were performed on the core with a range of the clay fraction
191 between 14-68% with a mean of 40%, and a range for the silt fraction between 30-71% and a mean
192 of 53%. Particle size fraction according to depth can be seen in Fig. 2c and the equivalent PSD in Fig.
193 3b.

4.2. Predictive models of particle size fraction and PSD

PLSR was used to generate quantitative models of the four main particle size fractions (clay, silt, very fine sand and fine sand) and also one to predict the 50 (out of 64) PSD classes from 0.15 µm to 259 µm of the core. The other classes do not have enough variabilities or spectral information to be modeled. Consequently, five models were created. The performance of each is presented in Table 2. All the models are validated with correlation coefficients (r_{cal} and r_{val}) and $r_{V/S}$ above 0.9 except for five classes between 6 µm and 10 µm of the PSD model with correlations greater than 0.5 (Supplementary material 3) due to very low variability in these classes. A similar number of selected wavelengths and LV was found between these models. Except for the silt fraction (4.0-62.5 µm) with a higher number of LV that induces a more complex spectral modeling, this agrees with the five classes between 6-10 µm that present lower modeling performances.

The selected wavelengths of these models allow us to identify the structural information that are correlated with the particle size effect (Viscarra Rossel and Behrens, 2010; Jacq et al., 2019b). The selected wavelengths for the grain size fractions and PSD predictions are consistent. This allows for identifying precisely the spectral areas of each particle size fraction. Nine spectral areas can be seen (Fig. 4):

- (1) Carotenoids and/or iron oxides compounds can be associated with wavelengths around 470 nm;
- (2,3) Two other areas can be linked to the presence of iron oxides presence (around 550 nm and 625 nm);
- (4) Chlorophyll pigments absorb around 700 nm;
- (5,7) The absorption wavelengths around 900 nm and 1450 nm can be due to water content, and this can be residual water that can be related to porosity;

- (6) Organic and hydroxyl bonds are associated with wavelengths around 1150 nm;
- (8) Mineralogical clay compounds wavelengths around 2200 nm are needed to predict physical clay fraction;
- (9) around 2400 nm are wavelengths associated with some mineral hydroxyl bonds.

Areas 1-5 and 7-8 are used for the prediction of fine fractions, such as clays and silts, whereas areas 5-6 and 9 are related to coarser fractions (very fine and fine sand).

Most of the selected wavelengths (34 over 44) are in the VNIR range, but that does not imply that the SWIR selected wavelengths are not relevant, because they contain relevant particle size information. Furthermore, the use of only one of the two sensors, independently of the other, does not allow to obtain a performant model. If only the VNIR is used, the performances of prediction are lower than 0.9, and the uncertainty is slightly higher. The same applies to the SWIR with performance lower than 0.9 and an inversion of the grain size of laminae (in which coarse grains are predicted as fine grains). This shows the importance of combining sensors of several spectral ranges to improve modeling performances with complementary information.

4.3. Prediction distribution

The predicted PSD in Fig. 3d evidence the same three units than the reference PSD in Fig. 3b (unit I: 22 µm, unit II: 22-6 µm and unit III: 6 µm). The down-sampling of the predicted PSD in Fig. 3c allows for surface and volume analysis to be linked together with a correlation of 0.96 ($p<0.05$) with the reference PSD. Two instantaneous deposits are quite visible at 14 cm and 16 cm depth through a marked drop in particle size.

The 50 models (one for each grain size class of PSD) allow estimation of the particle size fractions with the sum of the fine classes associated with the fractions according to the Wentworth scale

239 (Wentworth, 1922). The correlation between the prediction maps of a single or several models and
240 the bias (difference of the predicted values) have been estimated (Table 3, Supplementary material
241 4). This shows that the coarser the particle size, the lower the correlation. Moreover, there is more
242 bias for the silt fraction, which matches the observation of $r_{V/S}$ for the class between 6-10 µm. The $r_{V/S}$
243 is better for those estimated with a single model (Table 2), and the abundance maps (Supplementary
244 material 4) seem to be smooth for those estimated from the predicted PSD.

245 ## 4.4. Sediment sources

246 Five classes (EM) with their modes and abundances were estimated using the AnalySize function
247 applied to the PSD prediction (Fig. 5a-c):

248 1. The first one (EM1) has a mode at 0.42 µm (width at half height: 0.2-0.9 µm) and shows a
249 decreasing abundance trend from the bottom to the top of the core, from 25% to 10%.
250 2. The second (EM2) is centered at 4.19 µm (width at half height: 1.44-12.21 µm). It has an
251 upward decreasing trend from 62.5% to 5% and is mainly present below 25 cm depth and in
252 instantaneous event deposits with values higher than 30%.
253 3. The third (EM3) has a mode at 9.00 µm (width at half height: 4.88-15.21 µm) and has an
254 abundance quite stable at 10% except an increase at 15% in the instantaneous event.
255 4. The fourth (EM4) centered at 22.49 µm (width at half height: 12.21-45.51 µm) is mainly
256 present in the first 25 cm (eutrophic part) with an increasing upward trend from 5% to 30%.
257 5. The fifth (EM5) has a mode at 65.51 µm (width at half height: 26.2-150.0 µm) and is mainly
258 present in the first 25 cm (eutrophic part) with an increasing upward trend from 5% to 25%.
259 6. The residual part is mainly present in the two first units and has values lower than 2.5%.

260 The abundance maps in Fig. 6b-f show down and cross variations of the five pure classes. The two
261 first (EM1-2) have the same variations and a high correlation of 0.99 ($p < 0.05$). They characterize

unit III, i.e., the pre-eutrophication unit. Their trend along the core depth is the opposite of those of the two last (EM4-5) regardless of the proportions. EM4-5 are thus enriched in unit II and mostly unit I. The EM3 map is different with a quite homogeneous variation along the core except in specific lamina in the two first units. This end-member is also relatively abundant in the instantaneous event layer. The residual fraction (Fig. 6g) is more present in the laminated part of the core that might be explained by the presence of other sediment sources or a change in the size of previous sources.

5. Discussion

5.1. Interpretation of the sources based on the laminae succession

To interpret the end-members, we discuss a few laminae successions to see their compositions with the three main modes (EM2-3-4, Fig. 7) and to compare with previous descriptions of the laminae (Giguet-Covex et al., 2010) (Supplementary material 2). As presented in the lithological description, there are three different laminae, which are also well discriminated by the end-members (Fig. 7; Supplementary material 2. Based on PSD from the carbonated and decarbonated fractions of the sediments, Giguet-Covex et al. (2010) defined six classes corresponding to different sediment components. Five of these classes can be associated with the end-members determined in this study. EM4 agrees with the size range for large calcite crystals (15-30 µm) and diatoms (30-45 µm). EM3 reflecting small calcite crystals were characterized by a mode between 4.5 µm and 8 µm by Giguet-Covex et al. (2010). EM2 fits well with the detrital particles previously defined with a mode between 6 µm and 9 µm. EM1, which is highly correlated to EM2 (0.99 ($p < 0.05$)), was also observed but not discussed in the previous study (Giguet-Covex et al., 2010). This mode was only present in the carbonated PSD. We thus interpret it as representing micrite deposits. EM5 can be associated with a mix between diatom frustules and organic matter because it is highly correlated with EM4, which characterizes the spring lamina (i.e., the time of diatom bloom).

285 Due to the predicted PSD range (0.15 µm to 259 µm), two modes determined in the previous study
286 cannot be estimated with hyperspectral data (OM: 150-350 µm and very large carbonate particles:
287 160-300 µm). The modes are not the same due to the data used for their estimations. In Giguet-
288 Covex et al. (2010), PSD from the carbonated and decarbonated fractions of the sediments were
289 used. Whereas in our study, we use the complete PSD that corresponds to a mix between the two
290 fractions.

291 *5.1.1. Lamina counting*

292 The small calcite crystals (EM3; Fig. 7c), are quite localized and can be discriminated from the others.
293 It is possible to enhance it with image processing. An abundance threshold is fixed at 40% to binarize
294 the abundance map, in which upper values correspond to the EM3 deposit, and below this value to
295 another one. A median filter is used to remove solitary points. Then, the lamina can be easily
296 manually counted. They are correctly counted, until the year 1953, so 57 laminae can be
297 discriminated. Before this, they are too fuzzy and thin. Also, the small calcite crystal abundance
298 decreases lower than 40%, and even a linear decreasing threshold allows with difficulty to
299 discriminate this deposit because of laminae thinness.

300 *5.1.2.Estimation of the accumulation rate for each lamina*

301 The accumulation rate of each lamina can be estimated if the first pixel of each year is determined.
302 The EM2-3-4 maps were reduced to a median signal along the depth, the EM4 trend shows
303 fluctuations that can allow the beginning of the year to be identified. Unfortunately, it is too variable
304 to use automatic detection, so manual selections were made. Then, each pixel is associated with one
305 of the three laminae based on their EM percentage. Because the spring and summer laminae are
306 related to the same compounds (EM3-4) that overlap spatially, a rule based on the EM3-4 values is
307 proposed to separate them. If the value of the EM3 is greater than 40% of the EM4 one, then the

corresponding lamina is summer, otherwise, it is the spring. There is not this kind of problem with the winter lamina (EM2). Finally, this classification signal is used to count the number of pixels of each season in a year to estimate accumulation rates for each lamina (Fig. 7f and 8).

Fig. 8c shows an increase in diatom and large calcite crystal accumulations until 1982, followed by a decrease. These components reflect primary productivity. Consequently, a decrease after 1982 indicates an improvement in trophic status, which is in agreement with the decrease in phosphorus supplies presented in Jenny et al. (2013). The small calcite crystal accumulation rate has an opposite trend relative to the large calcite crystals and diatoms (Fig. 8b). This can also be explained by changes in trophic status. Eutrophication favors the precipitation of large calcite crystals compared to smaller ones. The accumulation rate of detrital particles (Fig. 8a) represents annual inputs from the Rhône River floods to the lake. The peak in 1990 represents the large floods of February 1990 (Jenny et al., 2014).

5.1.3.Visible and near-infrared spectra of each lamina

The three laminae, the instantaneous events and unit III were manually selected to find their pixels and then to extract their average spectral signatures (Fig. 9). The continuum removed (Clark and Roush, 1984) of the spectra allows for them to be normalized to compare the absorption spectra. It allows us to see five spectral ranges, which have different contributions in the different units and facies:

- The carotenoids and or iron oxides are mostly present in the laminated unit, in all the laminae.
- The chlorophyll pigments are mostly present in the spring and summer lamina and indicate the increase in lacustrine productivity due to eutrophication.

- The organic and hydroxyl compounds are higher in the summer and winter laminae as well as in the instantaneous events. This result is interpreted as reflecting lacustrine and terrestrial organic matter.
- The water content shows a gradient, from the lowest to the highest: non-eutrophic part, instantaneous event, winter, spring, summer laminae. This can be associated to residual water linked to porosity and compaction.

5.2. Detection of a sub-unit

The high resolution of analysis allows us to observe the trends with high precision and detect similar groups of trends for unit characterization. The Bourget core was divided into three units, but based on our evidence unit I was split into two sub-units, from 0 to 6 cm depth (2009-1992, unit I_a) and from 6 to 18 cm (1992-1962, unit I_b). The main observed changes in unit I are:

- A decrease of particle size in the PSD (Fig. 3d) that corresponds to an increase in small bio-induced calcite crystals, a decrease of large calcite crystals and diatom frustules (Fig. 5c and 6d-e),
- An increase in the residual part (Fig. 6g) that can be induced by modification of the environment.

The changes in this sub-unit can reflect the response of the lake to restoration programs, especially the diversion of treated waste-waters into the Rhône river in the 1990s that improved water quality and changed the biochemistry of the lake (Jacquet et al., 2005; Berthon et al., 2014).

5.3. Possibilities of the hyperspectral analysis

The use of high-resolution analysis, such as hyperspectral imaging, to estimate PSD associated with unmixing methods shows great possibilities to infer the provenance of the material and their

transport, as well as some deposition processes (floods, landslide, eruption). And also their variations across times, with changes in the population of diatoms or other organisms, changes in the physical-chemistry of the lake or the watershed or beyond.

Hyperspectral imaging is a very informative tool which can be used for more than particle size studies. This method has already proved useful for the prediction of organic matter (Van Exem et al., 2018; Jacq et al., 2019b), pigments (Schneider et al., 2018), and to perform source to sink analysis (Van Exem et al., 2019). Some proxies can be estimated directly after HSI acquisition, and other proxies need to process the data with or without destructive analysis.

A Matlab toolbox was created for the study of particle size presented in this paper and for a previous study of organic matter (Jacq et al., 2019b), and also allows for the generation of models for organic matter (LOI550 and RockEval) and particle size on several lakes (unpublished works). This toolbox can be found at https://github.com/JacqKevin/HSI_PLSR, developed with Matlab R2018a (the Parallel Computing Toolbox can be used for faster calculations). It contains functions to create a PLSR model and to apply it to the data, and also spectral preprocessing and variable selection algorithms. The regular use of this toolbox allows for highlighting precautions and difficulties related to data. The destructive data used to calibrate the models need to be representative of the main variations observed in the sample. The creation of an optimal predictive model requires to have variability inside the destructive analysis values. Future works will focus on the determination of relevant sampling areas for a proxy using only hyperspectral data. There are also the sediment matrix effects with the presence of several different lithologies that can disturb the modeling.

6. Conclusions

Hyperspectral imaging, with its spatial and spectral dimensions, is a very informative analytical method. This method promises great possibilities because it offers high resolution (200 µm), is non-

destructive and fast (15 min for a 1.5 m core for each sensor). In this study, HSI allows estimation of particle size distribution and size fraction in each point of the lake Bourget core. Then, with an unmixing method, five sediment sources were estimated and characterized in a previous study on this lake. The high spatial resolution allowed discrimination of three laminae. With image processing methods, they were discretized to estimate their spectral signatures and thus their chemistry. They were also manually counted, and that agreed with the age model. Therefore, hyperspectral imaging for the prediction of PSD and characterization of end-members, as well as their annual accumulation rates, is very useful for the study of sediment cores and environmental samples.

Some improvements are needed for the extraction of a relevant number of end-members in this data. Sequential extraction of end-members with the study of residuals seems to be relevant because the existing algorithms find only the main ones. Further works using the spatio-spectral approaches can improve the relevance of the estimated PSD, the fine particles can be estimated with the spectral dimension and the coarse ones with the spatial dimensions

Acknowledgements

The Bourget core used in this study is stored in the EDYTEM laboratory that also performed the PSD analysis in 2009 during the IPER-RETRO program (ANR-08-VUL 005). The hyperspectral imaging acquisition was made at the University of Normandie-Rouen and was funded by the Region Normandie, which supports the scientific consortium SCALE UMR CNRS 3730. Furthermore, we thank the anonymous reviewer and Jasper Knight (Editor-in-Chief) for their helpful comments.

References

Arnaud, F., Revel, M., Chapron, E., Desmet, M., Tribovillard, N., 2005. 7200 years of Rhône river flooding activity in Lake Le Bourget, France: a high-resolution sediment record of NW Alps hydrology. The Holocene 15, 420–428.

Arnaud, F., Révillon, S., Debret, M., Revel, M., Chapron, E., Jacob, J., Giguet-Covex, C., Poulenard, J., Magny, M., 2012. Lake Bourget regional erosion patterns reconstruction reveals Holocene NW European Alps soil evolution and paleohydrology. Quaternary Science Reviews 51, 81–92.

Berthon, V., Alric, B., Rimet, F., Perga, M.-E., 2014. Sensitivity and responses of diatoms to climate warming in lakes heavily influenced by humans. Freshwater Biology 59, 1755–1767.

Butz, C., Grosjean, M., Fischer, D., Wunderle, S., Tylmann, W., Rein, B., 2015. Hyperspectral imaging spectroscopy: a promising method for the biogeochemical analysis of lake sediments. Journal of Applied Remote Sensing 9, 1–20.

Butz, C., Grosjean, M., Goslar, T., Tylmann, W., 2017. Hyperspectral imaging of sedimentary bacterial pigments: a 1700-year history of meromixis from varved Lake Jaczno, northeast Poland. Journal of Paleolimnology 1–16.

Casa, R., Castaldi, F., Pascucci, S., Palombo, A., Pignatti, S., 2013. A comparison of sensor resolution and calibration strategies for soil texture estimation from hyperspectral remote sensing. Geoderma 197–198, 17–26.

Ce Liu, Yuen, J., Torralba, A., 2011. SIFT Flow: Dense Correspondence across Scenes and Its Applications. IEEE Transactions on Pattern Analysis and Machine Intelligence 33, 978–994.

Clark, R.N., Roush, T.L., 1984. Reflectance spectroscopy: Quantitative analysis techniques for remote

415 sensing applications. Journal of Geophysical Research: Solid Earth 89, 6329–6340.

416 de Santana, F.B., de Souza, A.M., Poppi, R.J., 2018. Visible and near infrared spectroscopy coupled to
417 random forest to quantify some soil quality parameters. Spectrochimica Acta - Part A:
418 Molecular and Biomolecular Spectroscopy 191, 454–462.

419 Debret, M., Chapron, E., Desmet, M., Rolland-Revel, M., Magand, O., Trentesaux, A., Bout-
420 Roumazeille, V., Nomade, J., Arnaud, F., 2010. North western Alps Holocene paleohydrology
421 recorded by flooding activity in Lake Le Bourget, France. Quaternary Science Reviews 29, 2185–
422 2200.

423 Demattê, J.A.M., Alves, M.R., Gallo, B.C., Fongaro, C.T., Souza, A.B., Romero, D.J., Sato, M.V., 2015.
424 Hyperspectral remote sensing as an alternative to estimate soil attributes. Revista Ciencia
425 Agronomica 46, 223–232.

426 Dietze, E., Hartmann, K., Diekmann, B., IJmker, J., Lehmkuhl, F., Opitz, S., Stauch, G., Wünnemann, B.,
427 Borchers, A., 2012. An end-member algorithm for deciphering modern detrital processes from
428 lake sediments of Lake Donggi Cona, NE Tibetan Plateau, China. Sedimentary Geology 243–244,
429 169–180.

430 Giguet-Covex, C., Arnaud, F., Poulenard, J., Enters, D., Reyss, J.-L., Millet, L., Lazzaroto, J., Vidal, O.,
431 2010. Sedimentological and geochemical records of past trophic state and hypolimnetic anoxia
432 in large, hard-water Lake Bourget, French Alps. Journal of Paleolimnology 43, 171–190.

433 Hermansen, C., Knadel, M., Moldrup, P., Greve, M.H., Karup, D., de Jonge, L.W., 2017. Complete Soil
434 Texture is Accurately Predicted by Visible Near-Infrared Spectroscopy. Soil Science Society of
435 America Journal 81, 758–769.

436 Heslop, D., von Dobeneck, T., Höcker, M., 2007. Using non-negative matrix factorization in the

437 “unmixing” of diffuse reflectance spectra. Marine Geology 241, 63–78.

438 Jacq, K., Coquin, D., Fanget, B., Perrette, Y., Debret, M., 2019a. Study of pansharpening methods
439 applied to hyperspectral images of sediment cores, in: 2019 22nd International Conference on
440 Information Fusion (FUSION) (FUSION 2019). Ottawa, Canada.

441 Jacq, K., Perrette, Y., Fanget, B., Sabatier, P., Coquin, D., Martinez-Lamas, R., Debret, M., Arnaud, F.,
442 2019b. High-resolution prediction of organic matter concentration with hyperspectral imaging
443 on a sediment core. Science of the Total Environment 663, 236–244.

444 Jacquet, S., Briand, J.-F., Leboulanger, C., Avois-Jacquet, C., Oberhaus, L., Tassin, B., Vinçon-Leite, B.,
445 Paolini, G., Druart, J.-C., Anneville, O., Humbert, J.-F., 2005. The proliferation of the toxic
446 cyanobacterium Planktothrix rubescens following restoration of the largest natural French lake
447 (Lac du Bourget). Harmful Algae 4, 651–672.

448 Jenny, J.-P., 2013. Réponses des grands lacs périalpins aux pressions anthropiques et climatiques
449 récentes : reconstitutions spatio-temporelles à partir d’archives sédimentaires. HAL. Université
450 Grenoble Alpes.

451 Jenny, J.-P., Arnaud, F., Dorioz, J.-M., Covex, C.G., Frossard, V., Sabatier, P., Millet, L., Reyss, J.-L.,
452 Tachikawa, K., Bard, E., Pignol, C., Soufi, F., Romeyer, O., Perga, M.-E., 2013. A spatiotemporal
453 investigation of varved sediments highlights the dynamics of hypolimnetic hypoxia in a large
454 hard-water lake over the last 150 years. Limnology and Oceanography 58, 1395–1408.

455 Jenny, J.-P., Wilhelm, B., Arnaud, F., Sabatier, P., Giguet Covex, C., Mélo, A., Fanget, B., Malet, E.,
456 Ployon, E., Perga, M.E., 2014. A 4D sedimentological approach to reconstructing the flood
457 frequency and intensity of the Rhône River (Lake Bourget, NW European Alps). Journal of
458 Paleolimnology 51, 469–483.

459 Nanni, M.R., Cezar, E., Silva Junior, C.A. da, Silva, G.F.C., da Silva Gualberto, A.A., 2018. Partial least
460 squares regression (PLSR) associated with spectral response to predict soil attributes in
461 transitional lithologies. Archives of Agronomy and Soil Science 64, 682–695.

462 Nouri, M., Gomez, C., Gorretta, N., Roger, J.M., 2017. Clay content mapping from airborne
463 hyperspectral Vis-NIR data by transferring a laboratory regression model. Geoderma 298, 54–
464 66.

465 Paterson, G.A., Heslop, D., 2015. New methods for unmixing sediment grain size data. Geochemistry,
466 Geophysics, Geosystems 16, 4494–4506.

467 Perga, M.-E., Frossard, V., Jenny, J.-P., Alric, B., Arnaud, F., Berthon, V., Black, J.L., Domaizon, I.,
468 Giguet-Covex, C., Kirkham, A., Magny, M., Manca, M., Marchetto, A., Millet, L., Paillès, C.,
469 Pignol, C., Poulenard, J., Reyss, J.-L., Rimet, F., Sabatier, P., Savichtcheva, O., Sylvestre, F.,
470 Verneaux, V., 2015. High-resolution paleolimnology opens new management perspectives for
471 lakes adaptation to climate warming. Frontiers in Ecology and Evolution 3, 72.

472 Van Exem, A., Debret, M., Copard, Y., Vannière, B., Sabatier, P., Marcotte, S., Laignel, B., Reyss, J.-L.,
473 Desmet, M., 2018. Hyperspectral core logging for fire reconstruction studies. Journal of
474 Paleolimnology 59, 297–308.

475 van Hateren, J.A., Prins, M.A., van Balen, R.T., 2017. On the genetically meaningful decomposition of
476 grain-size distributions: A comparison of different end-member modelling algorithms.
477 Sedimentary Geology 375, 49–71.

478 Viscarra Rossel, R.A., Behrens, T., 2010. Using data mining to model and interpret soil diffuse
479 reflectance spectra. Geoderma 158, 46–54.

480 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., Brodsky, L., Du, C.W., Chappell, A., Fouad, Y., Genot, V., Gomez, C., Grunwald, S., Gubler, A., Guerrero, C., Hedley, C.B., Knadel, M., Morràs, 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., Ji, W., 2016. A global spectral library to characterize the world's soil. Earth-Science Reviews.

Wentworth, C.K., 1922. A Scale of Grade and Class Terms for Clastic Sediments. The Journal of Geology 30, 377–392.

Wetzel, D.L., 1983. Near-infrared reflectance analysis. Analytical Chemistry 55, 1165A-1176A.

Wold, S., Ruhe, A., Wold, H., Dunn, III, W.J., 1984. The Collinearity Problem in Linear Regression. The Partial Least Squares (PLS) Approach to Generalized Inverses. SIAM Journal on Scientific and Statistical Computing 5, 735–743.

Yu, S.-Y., Colman, S.M., Li, L., 2016. BEMMA: A Hierarchical Bayesian End-Member Modeling Analysis of Sediment Grain-Size Distributions. Mathematical Geosciences 48, 723–741.

Żarczyński, M., Szmańda, J., Tylmann, W., Żarczyński, M., Szmańda, J., Tylmann, W., 2019. Grain-Size Distribution and Structural Characteristics of Varved Sediments from Lake Żabińskie (Northeastern Poland). Quaternary 2, 1–15.

Zhang, X., Zhou, A., Wang, X., Song, M., Zhao, Y., Xie, H., Russell, J.M., Chen, F., 2017. Unmixing grain-size distributions in lake sediments: a new method of endmember modeling using hierarchical clustering. Quaternary Research 1–9.

Figures

Fig. 1. (a) Location and the catchment area of Lake Bourget. (b) Bathymetry, tributaries and effluents of the lake with the coring site.

Fig. 2. (a) Image RGB of the core, (b) lithology (with three units and flood deposits), (c) particle size fractions (in %) by the laser granulometer and (d) average predicted particle size fractions (in %) by the hyperspectral imaging.

Fig. 3. (a) Image RGB of the core, (b) PSD (in %) of the laser granulometer, (c) predicted PSD (in %) subsampled at the sampling resolution (5 mm), and (d) high-resolution predicted PSD (in %) by the hyperspectral imaging.

Fig. 4. Selected wavelengths and associated compound areas for the five models (four fractions and the complete PSD) in the visible range on the left, the near-infrared range on the right and a gap between them in the grey area.

Fig. 5. (a) The five end-members estimated with the predicted PSD by HSI, (b) image RGB of the core, (c) profiles of the five sources (1-5) and the residual part (6) estimated with the averaged of the EM maps along depth (in %).

Fig. 6. (a) Image RGB of the core, (b-f) the abundance maps of the five end-members estimated by AnalySize and (g) the residual fraction (in %).

Fig. 7. (a) Image RGB of a laminated area, (b-d) the three main end-members abundance maps (EM4-3-2 in %), (e) schematic succession of the lamina; (f) estimation of the annual accumulation source rates (in mm/year).

Fig. 8. Estimation of the annual lamina accumulation rates for (a) detrital particles, (b) small calcite crystals, (c) diatoms frustules associated with large calcite crystals.

Fig. 9. (a) Visible and near-infrared signatures and (b) continuum removal with associated main absorptions of the three laminae (spring, summer, winter), an instantaneous event and the non-eutrophic part.

Tables

Table 1. Hyperspectral sensor properties

Table 2. Statistics of the reference data and models performances (number of latent variables (LV) and of selected wavelengths (wl), calibration and validation correlations (r_{cal} and r_{val}), uncertainty (Unc), volume-surface correlation ($r_{V/S}$)) for the four particle size fractions and the averaged of the fifty classes of the PSD.

Table 3. Correlation and bias between the abundance maps of the particle size fractions estimated by their PLSR models or with the PSD prediction; volume-surface correlation of the map estimated with the PSD prediction.

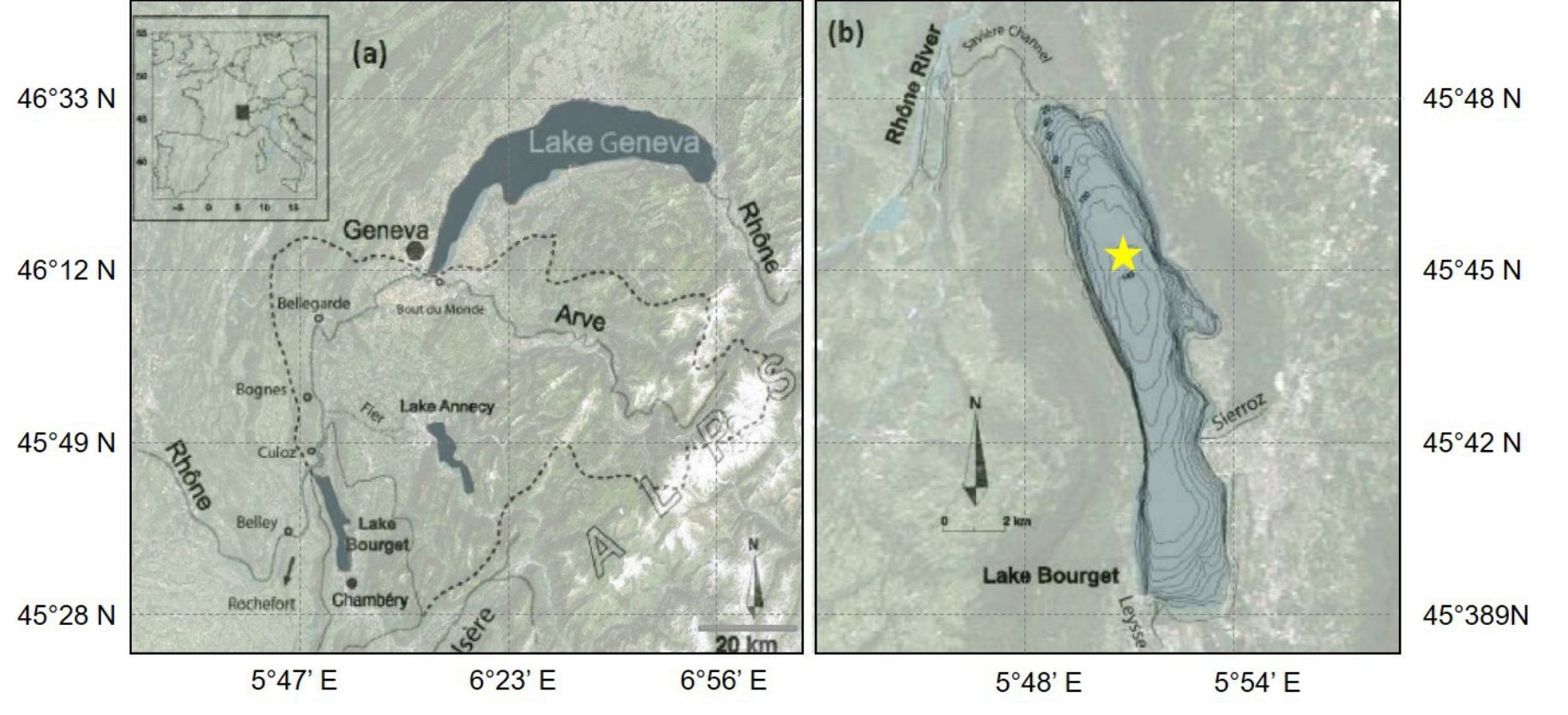
Supplementary materials

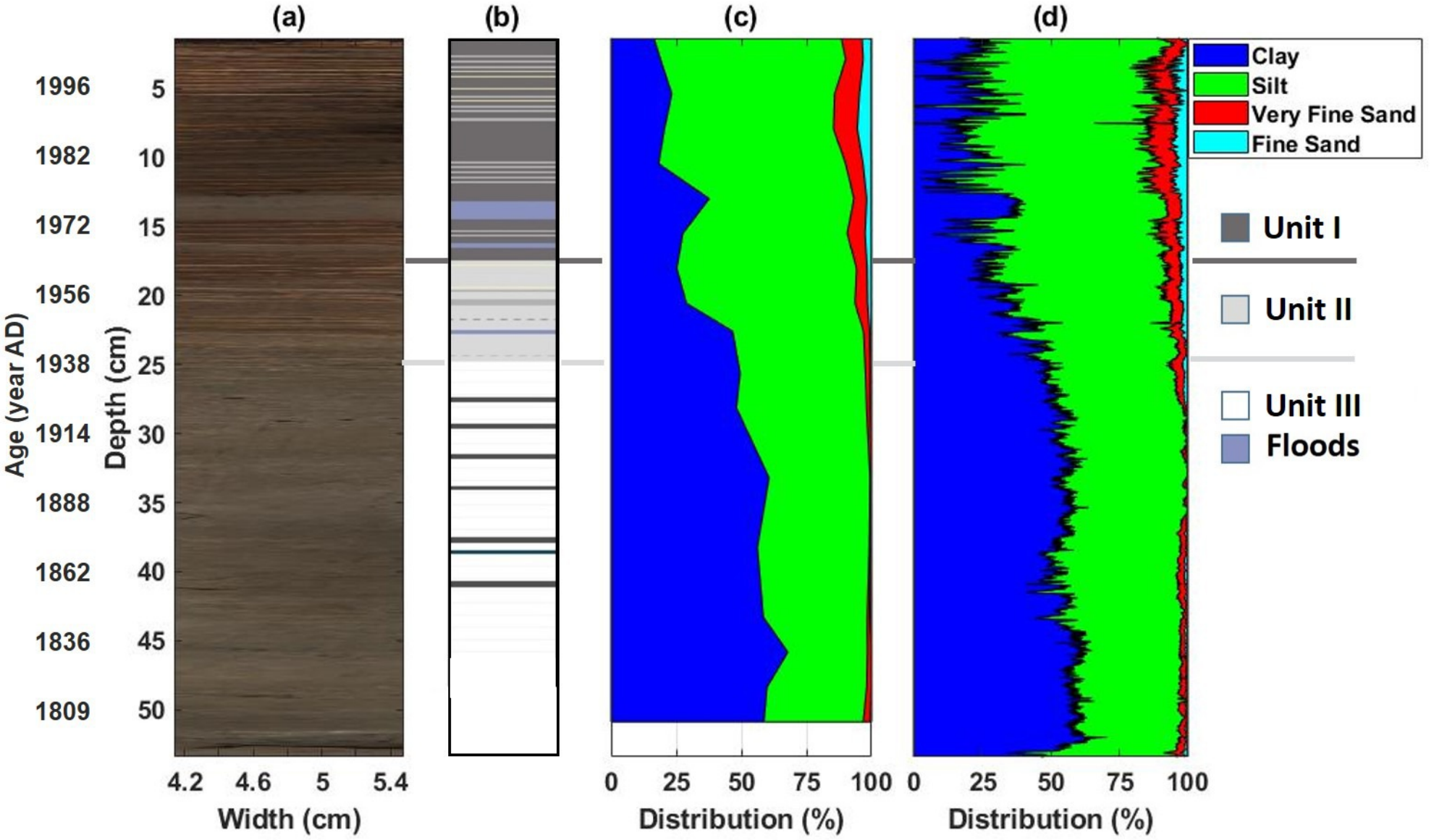
Supplementary material 1. Correlations of the sediment cores use for our (LDB09_P101) and Giguet (LDB06_P2) studies *(Giguet-Covex et al., 2010)*

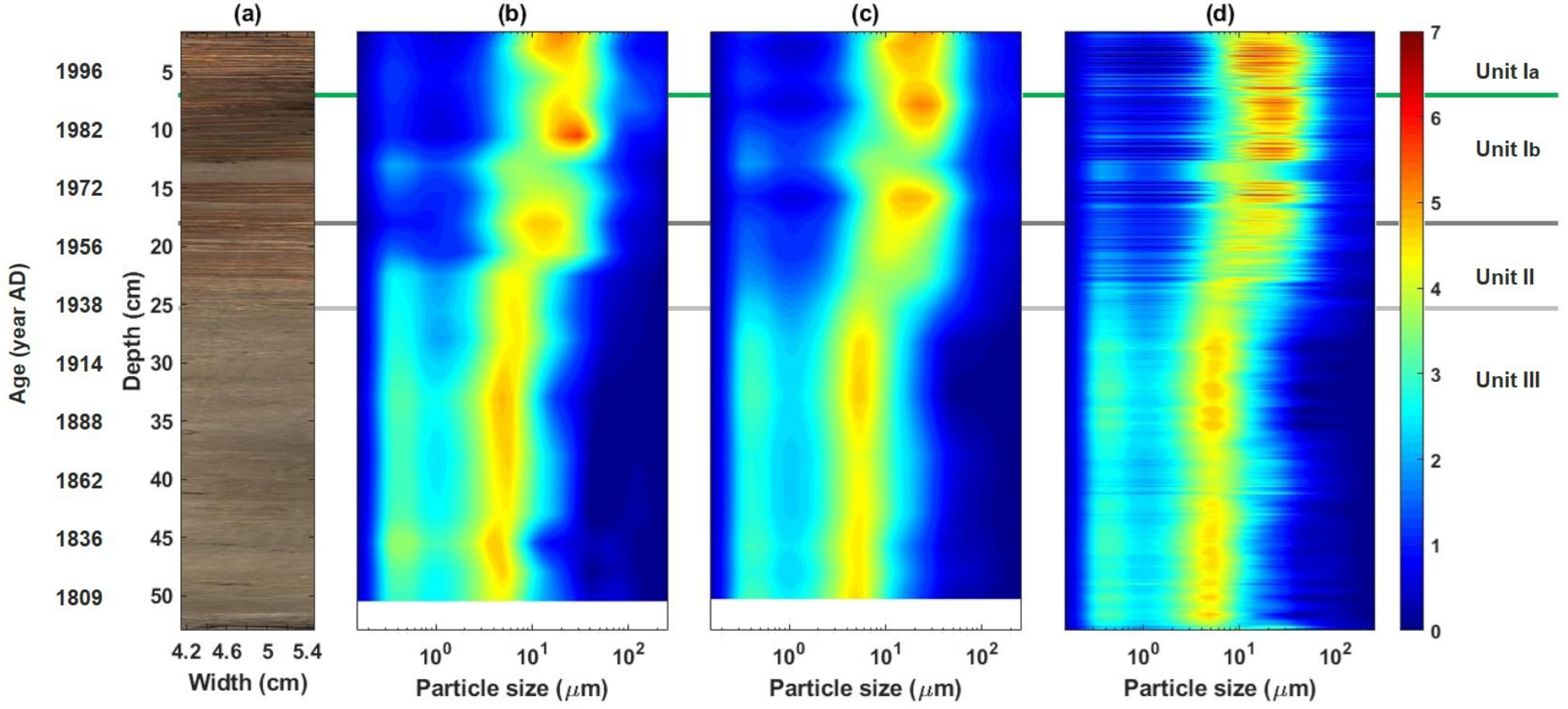
Supplementary material 2. Thin section of the LDB06_P2 core used for the Ca, Al and Si counts by micro-XRF.

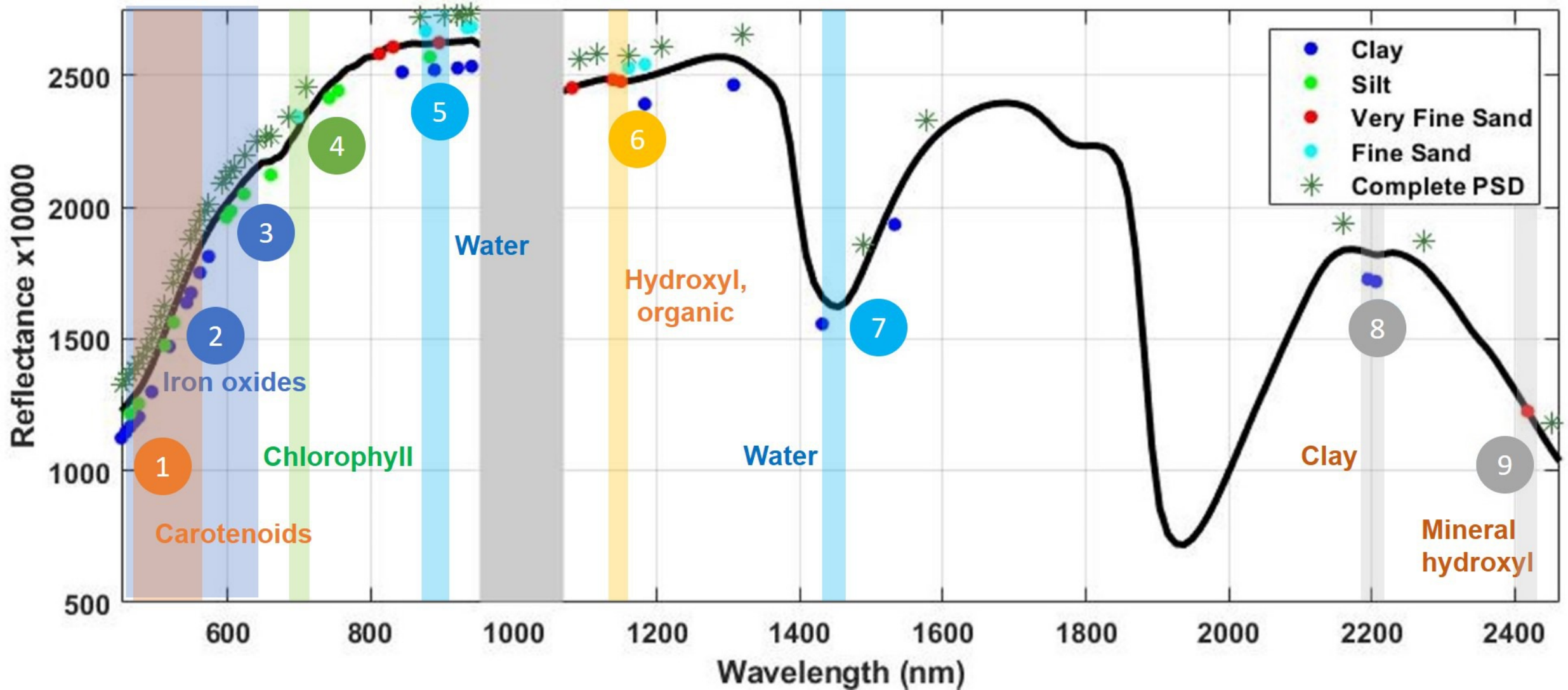
Supplementary material 3. Correlation coefficients of the PSD models in calibration, validation and for $r_{V/S.}$

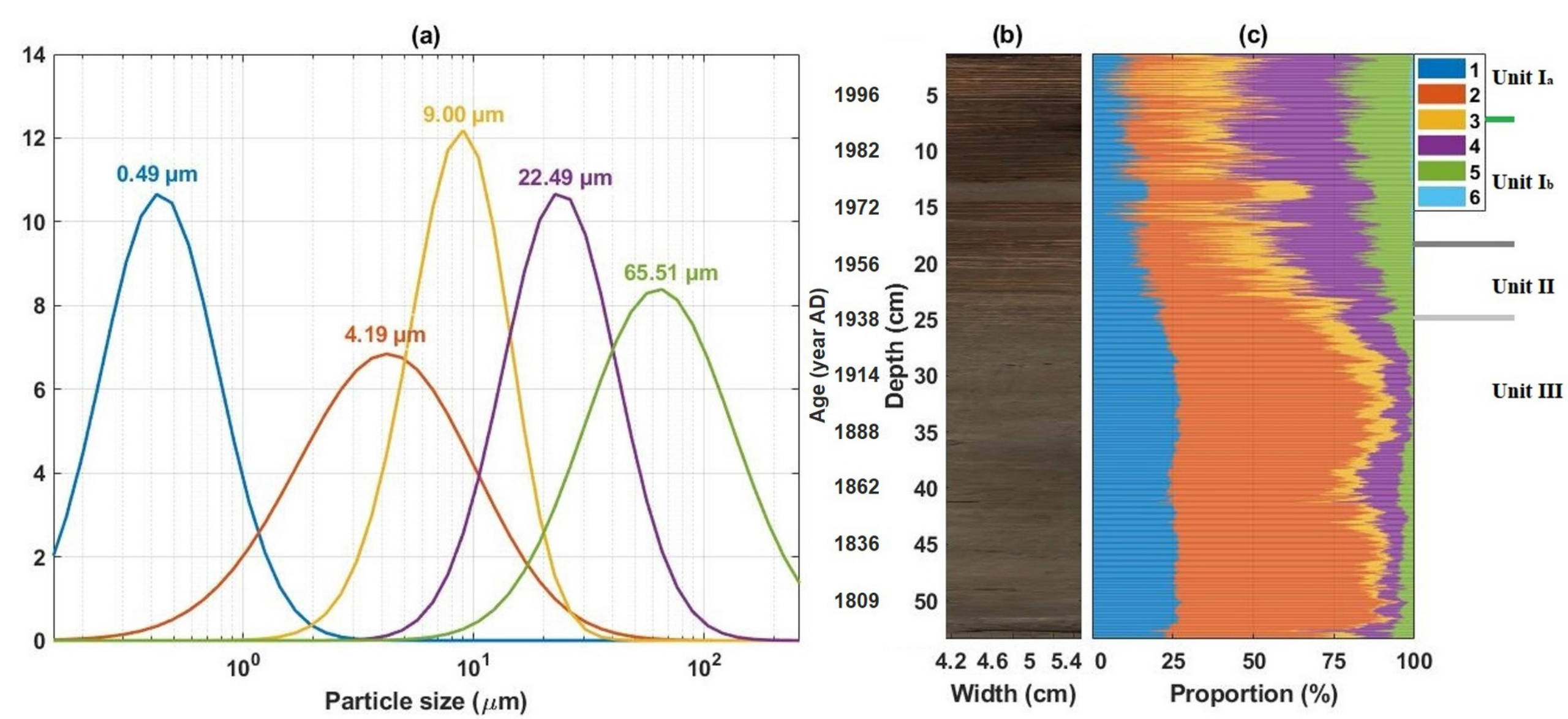
Supplementary material 4. Abundance maps of the particle size fractions estimated with their models (single, a-c-e-g) or with the predicted PSD (multiple, b-d-f-h) (in %).

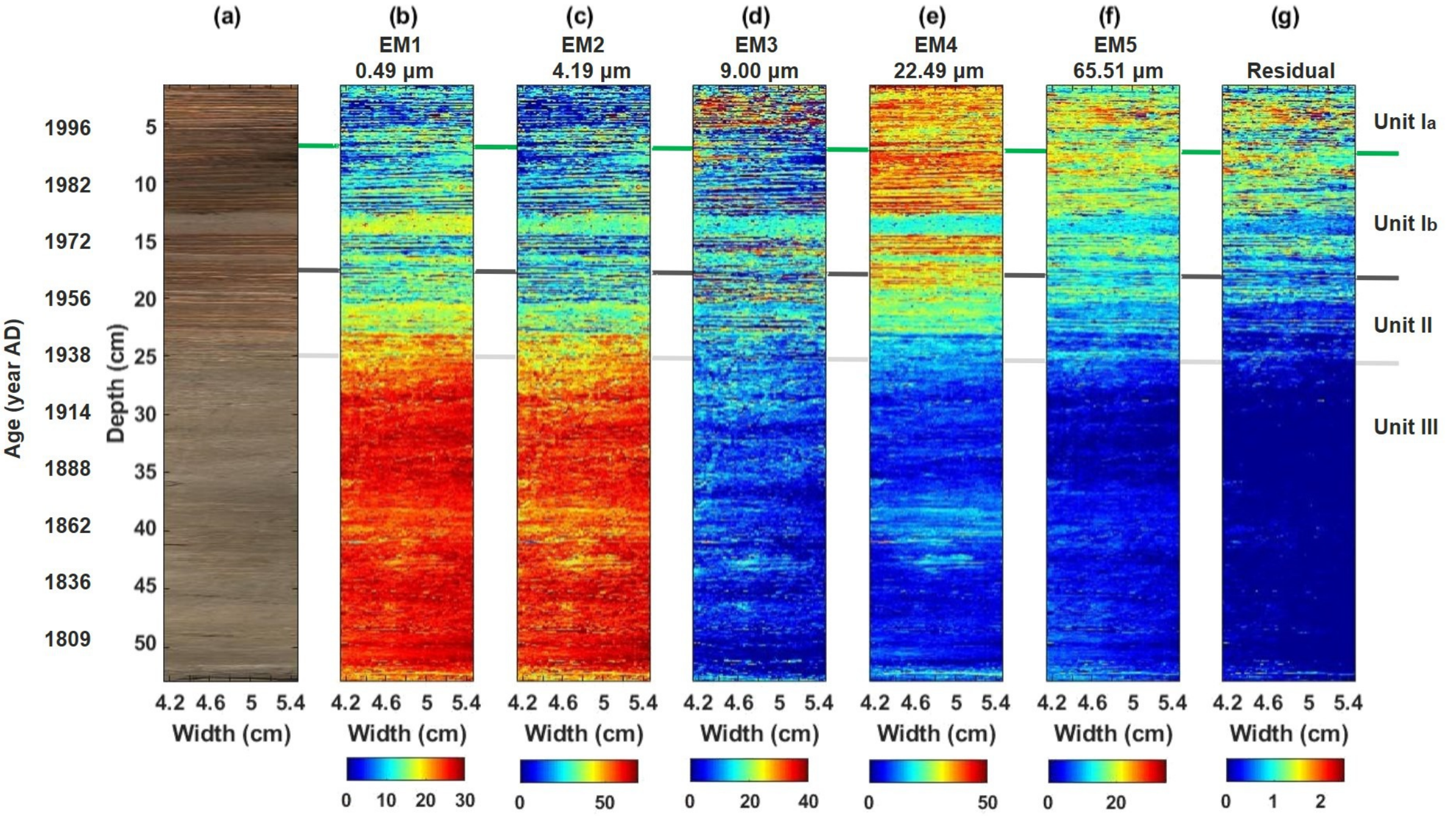

(a)
Lake Geneva
Geneva
Rhône
Bellegarde
Bout du Monde
Arve
Bognes
Fier
Lake Annecy
Culoz
Rhône
Belley
Lake Bourget
Chambéry
Rochefort
ALPS
N
20 km
(b)
Rhône River
Savière Channel
Sierroz
N
0
2 km
Lake Bourget
Leysse
46°33 N
46°12 N
45°49 N
45°28 N
45°48 N
45°45 N
45°42 N
45°389N
5°47’ E
6°23’ E
6°56’ E
5°48’ E
5°54’ E

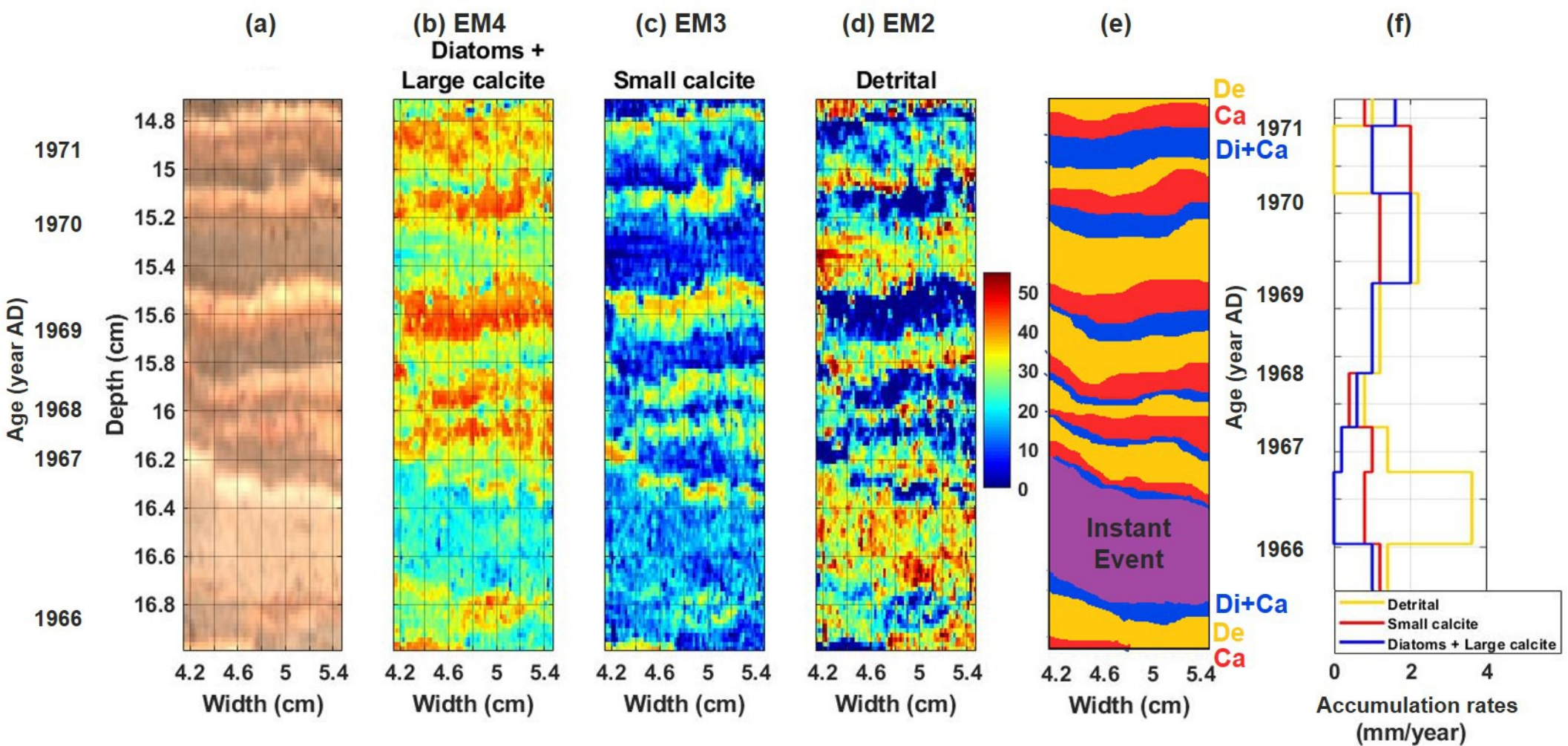

(a)
(b)
(c)
(d)
Clay
Silt
Very Fine Sand
Fine Sand
Unit I
Unit II
Unit III
Floods
Age (year AD)
1996
1982
1972
1956
1938
1914
1888
1862
1836
1809
Depth (cm)
5
10
15
20
25
30
35
40
45
50
4.2
4.6
5
5.4
Width (cm)
0
25
50
75
100
Distribution (%)
0
25
50
75
100
Distribution (%)

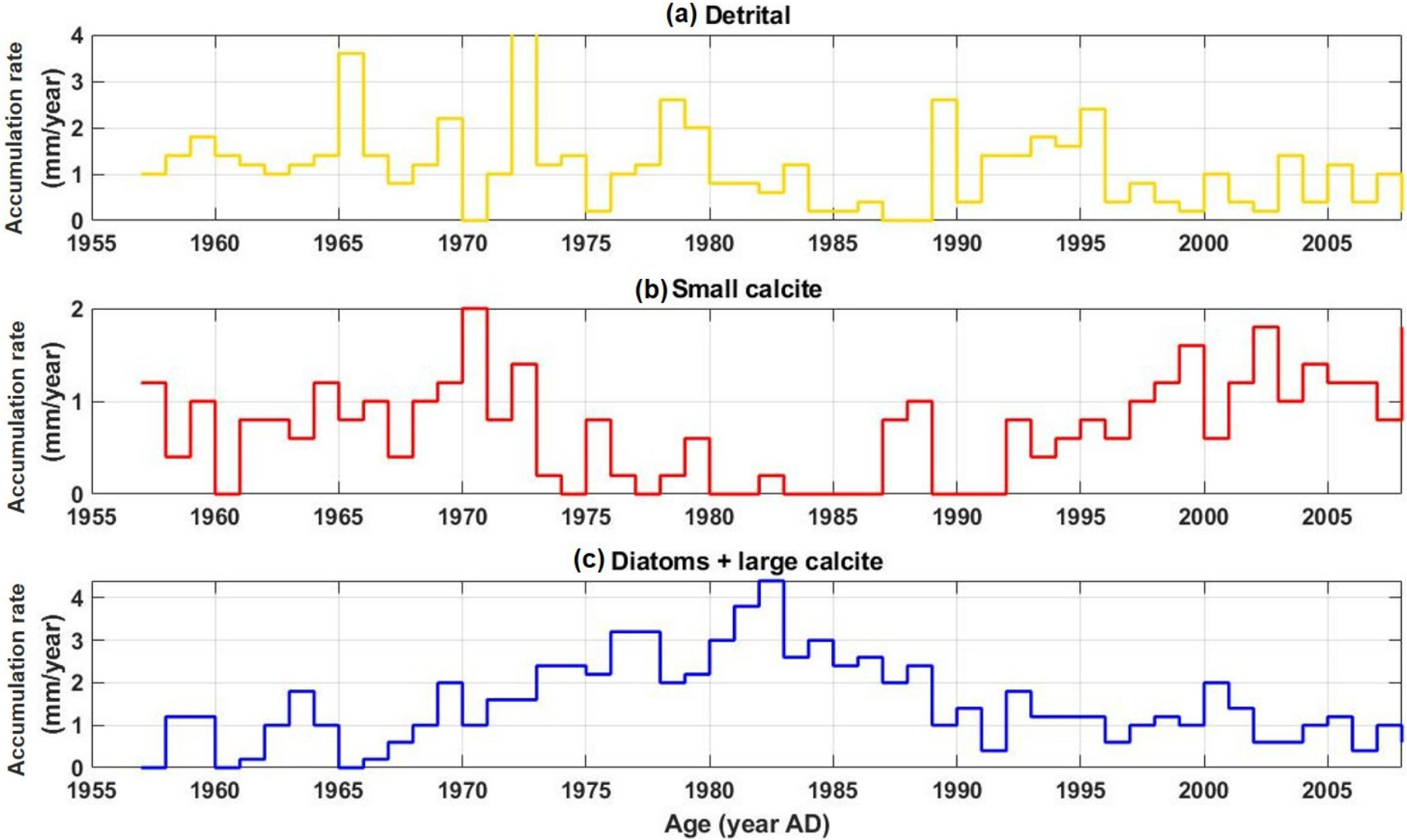

(a)
(b)
(c)
(d)
Age (year AD)
1996
1982
1972
1956
1938
1914
1888
1862
1836
1809
Depth (cm)
5
10
15
20
25
30
35
40
45
50
4.2 4.6 5 5.4
Width (cm)
10^0 10^1 10^2
Particle size (μm)
10^0 10^1 10^2
Particle size (μm)
10^0 10^1 10^2
Particle size (μm)
7
6
5
4
3
2
1
0
Unit Ia
Unit Ib
Unit II
Unit III

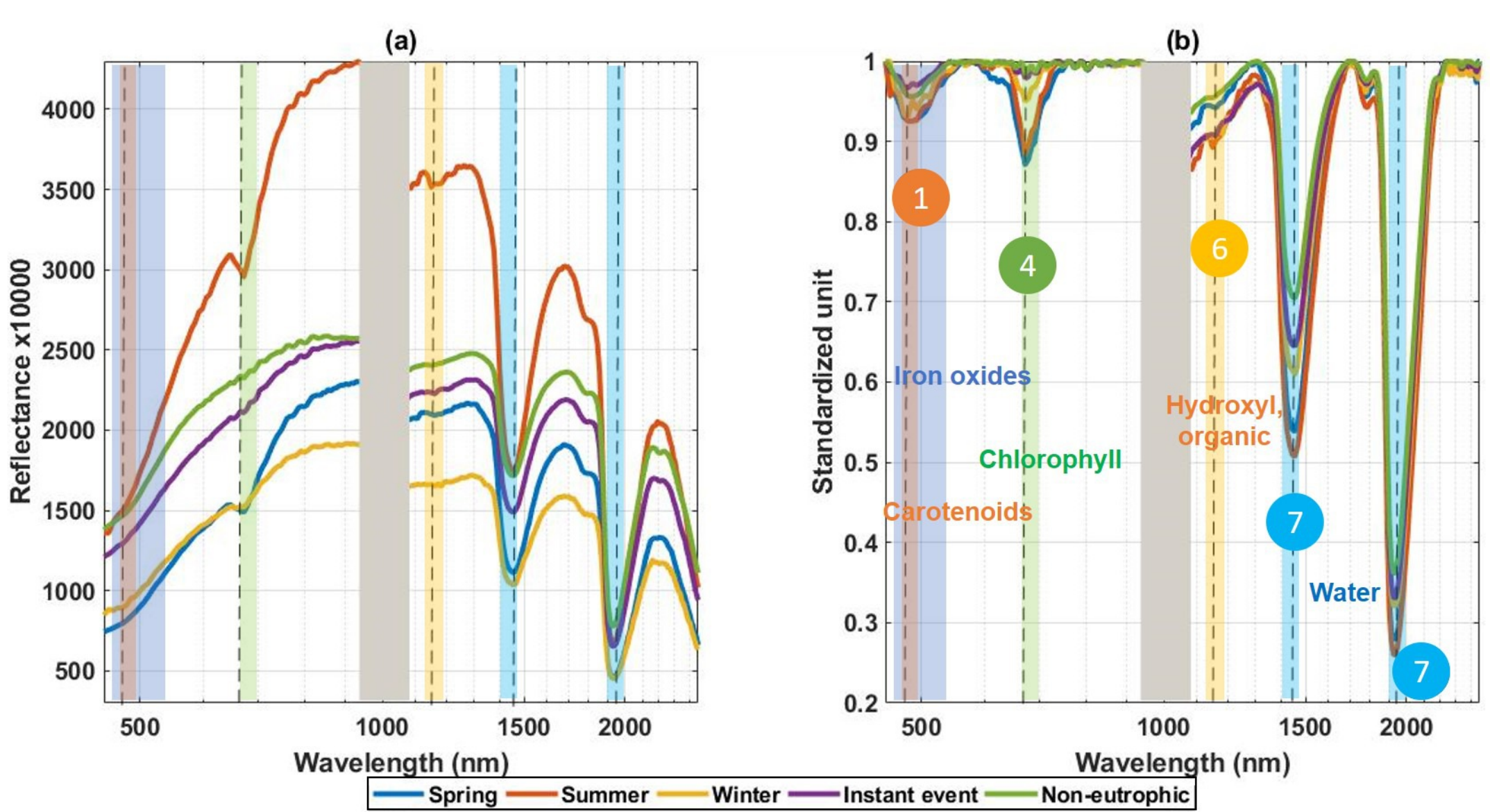

Clay
Silt
Very Fine Sand
Fine Sand
Complete PSD
Reflectance x10000
Wavelength (nm)
2500
2000
1500
1000
500
600
800
1000
1200
1400
1600
1800
2000
2200
2400
1
2
3
4
5
6
7
8
9
Carotenoids
Iron oxides
Chlorophyll
Water
Hydroxyl,
organic
Water
Clay
Mineral
hydroxyl


(a)
(b)
(c)
0.49 μm
4.19 μm
9.00 μm
22.49 μm
65.51 μm
Particle size (μm)
Age (year AD)
Depth (cm)
Width (cm)
Proportion (%)
1
2
3
4
5
6
Unit Ia
Unit Ib
Unit II
Unit III


(a)
(b)
EM1
0.49 μm
(c)
EM2
4.19 μm
(d)
EM3
9.00 μm
(e)
EM4
22.49 μm
(f)
EM5
65.51 μm
(g)
Residual
Age (year AD)
1996
1982
1972
1956
1938
1914
1888
1862
1836
1809
Depth (cm)
5
10
15
20
25
30
35
40
45
50
Unit Ia
Unit Ib
Unit II
Unit III
4.2 4.6 5 5.4
Width (cm)
0 10 20 30
0 50
0 20 40
0 50
0 20
0 1 2


(a)
(b) EM4
Diatoms + Large calcite
(c) EM3
Small calcite
(d) EM2
Detrital
(e)
(f)
Age (year AD)
1971
1970
1969
1968
1967
1966
Depth (cm)
14.8
15
15.2
15.4
15.6
15.8
16
16.2
16.4
16.6
16.8
4.2 4.6 5 5.4
Width (cm)
50
40
30
20
10
0
De
Ca
Di+Ca
Instant Event
Di+Ca
De
Ca
Age (year AD)
Detrital
Small calcite
Diatoms + Large calcite
0 2 4
Accumulation rates (mm/year)


(a) Detrital
Accumulation rate (mm/year)
0
1
2
3
4
1955
1960
1965
1970
1975
1980
1985
1990
1995
2000
2005
(b) Small calcite
Accumulation rate (mm/year)
0
1
2
(c) Diatoms + large calcite
Accumulation rate (mm/year)
Age (year AD)


(a)
Reflectance x10000
4000
3500
3000
2500
2000
1500
1000
500
500
1000
1500
2000
Wavelength (nm)
(b)
Standardized unit
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
1
4
6
7
7
Iron oxides
Chlorophyll
Carotenoids
Hydroxyl, organic
Water
500
1000
1500
2000
Wavelength (nm)
Spring
Summer
Winter
Instant event
Non-eutrophic

Name	Spectral range (nm)	Spectral resolution	Spatial resolution (μm)
Visible and Near InfraRed (VNIR)	400-1000	6	50
Short-Wave InfraRed (SWIR)	968-2574	12	200

		Reference			Prediction					
	n	**Range (%)**	**Mean (%)**	**Std (%)**	**LV**	**wl**	r_{cal}	r_{val}	**Unc. (%)**	$r_{V/S}$
Clay (1-3.9 µm)	18	14.33-67.85	39.94	18.06	4	21	0.91	0.94	7.39	0.96
Silt (3.9-62.5 µm)		30.49-71.09	53.56	13.40	7	11	0.91	0.90	9.01	0.92
Very Fine Sand (62.5-125 µm)		0.44-9.33	4.11	3.06	4	7	0.86	0.95	1.32	0.89
Fine Sand (125-250 µm)		0.05-5.36	1.77	1.63	4	6	0.93	0.95	0.86	0.69
50 PSD classes (averaged values)		0.68-3.40	1.99	0.90	4	44	0.85	0.83	0.71	0.93

Fraction	Correlation (p-value)	Bias (%)	$r_{V/S}$
Clay (1-3.9 µm)	0.95 (<0.05)	2.14	0.91
Silt (3.9-62.5 µm)	0.61 (<0.05)	9.17	0.93
Very Fine Sand (62.5-125 µm)	0.94 (<0.05)	0.03	0.77
Fine Sand (125-250 µm)	0.92 (<0.05)	-0.20	0.72