



HAL
open science

Improvement of the chemical content prediction of a model powder system by reducing multiple scattering using polarized light spectroscopy

R. Bendoula, Alexia Gobrecht, Bruno Moulin, Jean-Michel Roger, Véronique Bellon-Maurel

► To cite this version:

R. Bendoula, Alexia Gobrecht, Bruno Moulin, Jean-Michel Roger, Véronique Bellon-Maurel. Improvement of the chemical content prediction of a model powder system by reducing multiple scattering using polarized light spectroscopy. *Applied Spectroscopy*, 2015, 69 (1), pp.95-102. 10.1366/14-07539 . hal-01152872

HAL Id: hal-01152872

<https://hal.science/hal-01152872>

Submitted on 18 May 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Improvement of the chemical content prediction of a model powder system by reducing multiple scattering using polarized light spectroscopy

Ryad Bendoula,^{1,*} Alexia Gobrecht,¹ Benoit Moulin,¹ Jean-Michel Roger¹

and Veronique Bellon-Maurel¹

¹*Irstea, UMR ITAP, 361 rue J-F Breton, F-34196 Montpellier, France*

*ryad.bendoula@irstea.fr

Abstract: Near Infrared Spectroscopy (NIRS) is a powerful non destructive analytical method used to analyze major compounds in bulk materials and products and requiring no sample preparation. It is widely used in routine analysis and also in-line in industries, in-vivo with biomedical applications or in-field for agricultural and environmental applications. However, highly scattering samples subvert Beer-Lambert law's linear relationship between spectral absorbances and the concentrations. Instead of spectral pre-processing, which is commonly used by NIR spectroscopists to mitigate the scattering effect, we put forward an optical method, i.e. coupling polarized light with NIR spectrometry, to free spectra from scattering effect. This should allow us to retrieve linear and steady conditions for spectral analysis. When tested in Visible-NIR (VIS-NIR) range (400-800 nm) on model media - mixtures of scattering and absorbing particles - the set-up provided

21 significant improvements in absorber concentration estimation precision as well as in
22 the quality and robustness of the calibration model.

23 **Index Heading:** Polarized Light Spectroscopy, Multiple Light Scattering, Diffuse
24 Reflectance, Optical Properties, Chemical Content Prediction.

25 **1. Introduction**

26 Visible - Near infrared spectroscopy (VIS-NIRS) is a well-known technique used for
27 measuring the chemical composition of a wide variety of media and products.

28 Although VIS-NIRS has been quoted in articles for approximately 50 years^{1,2} with
29 this purpose, it really took off in the late 80's in agricultural and food applications
30 (jumping from around 10 publications per year in the late 80's to 150 publications per
31 year in the turn of century), and then in the 90's for pharmaceutical and biomedical
32 applications. Today, it plays a major role in these sectors, as a routine laboratory
33 method for in vivo or in-line monitoring system. On the one hand VIS-NIRS presents
34 several advantages: VIS-NIR extinction coefficients are small compared to mid-
35 infrared (MIR) ones, which allows light to penetrate deeper into objects and avoids
36 time-consuming sample preparation; VIS-NIR light scattering makes it possible to
37 analyze bulk samples with a retro-diffusion optical configuration, thus turning it into
38 a non-destructive technique. In addition VIS-NIR optical components are low cost
39 and with high Signal-to-Noise Ratio (SNR). On the other hand, VIS-NIRS has
40 several drawbacks: the VIS-NIR spectrum is poorly resolved as it is made up of
41 scattering effects and of wide low-intensity harmonics and combinations of MIR

42 fundamental absorption bands. Consequently, retrieving chemical information from
43 VIS-NIR spectra is quite painstaking and requires advanced chemometrics: it is
44 based on calibration models to be built between VIS-NIR spectra and known
45 concentrations of a set of calibration samples. Traditionally, linear multivariate
46 calibration methods such as Principal Component Regression (PCR) and Partial
47 Least Square Regression (PLSR) are used in VIS-NIRS. However, scattering effects
48 are troublesome in the VIS-NIR spectra of turbid media, defined by Shi and
49 Anderson³ as "*samples exhibiting multiple scattering events*". Scattering can be
50 several orders of magnitude larger than absorption⁴ and may invalidate the use of
51 such data processing methods, which are themselves based on the underlying
52 assumption of a linear Beer-Lambert law relationship between absorbance spectra
53 and chemical concentration⁵. It is therefore necessary for VIS-NIR spectroscopists
54 working on highly scattering media to use strategies to release VIS-NIR spectra from
55 scattering effects. The most common strategy is spectral pre-treatment. These
56 preprocessing step is specifically designed to reduce multiplicative and additive
57 effects caused by variations in sample physical properties^{6,7}. Among them, standard
58 normal variate (SNV) often associated to detrend⁸, multiplicative signal correction
59 (MSC)⁹, Extended MSC (EMSC)⁷, normalization or Optical Path Length Estimation
60 and Correction (OPLEC)^{5,10}. However, these approaches remain questionable: they
61 consider that scattering is nearly constant over the wavelengths, which is not the
62 case³; they may eliminate chemical-related information, which is very small with

63 regard to scattering effects¹¹; they are inappropriate when light scattering varies
64 greatly from sample to sample¹².

65 Another option is to acquire the spectrum in a way that separates the part related to
66 absorption from the part related to scattering. Specific experimental techniques,
67 related to the application of light propagation theory and resolution of the Radiative
68 Transfer Equation³ have been proposed, including adding-doubling set-ups¹²⁻¹⁴,
69 spatially-resolved spectroscopy¹⁵, time-resolved spectroscopy^{4,16} and frequency-
70 resolved spectroscopy¹⁷.

71 Although powerful, these methods have their limitations, particularly when applied
72 on highly turbid samples. First, they may require complex and sometimes expensive
73 optical implementations, which may not be compatible with conventional
74 spectrometers or with highly turbid samples (for which transmission measurement is
75 not possible). Secondly, as they rely on the estimation of absorption and scattering
76 coefficients achieved by model inversion, parameters describing the studied medium
77 (sample thickness, refractive index, particle size and shape...) must be known or
78 approximated, which may be a troublesome task as they are often unknown in
79 complex media^{12,18}.

80 Whereas separating absorption and scattering from a VIS-NIR signal is still an open
81 research issue on highly turbid samples, the main demand from VIS-NIR
82 spectroscopists is *merely* for spectra with reduced impact of scattering in order to
83 better fit Beer-Lambert's Law conditions^{19,20}.

84 Light polarization subtraction is a simple technique to reduce directly the effects of
85 multi-scattering on the measured signal^{20,21}. This approach has been based on the fact
86 that, when light interacts with matter, a small number of scattering events do not
87 significantly modify the polarization status of the beam whereas multiple scattering
88 leads to depolarization^{18,22}. Polarization subtraction technique^{19,23-25} was used to
89 select light beams that retain initial polarization and which are therefore less
90 impacted by multiple scattering events.

91 Although this technique has gained interest in the field of biomedical^{20,21,25}, where it
92 is used to optically target subsurface organelles (particles suspended in water) and
93 tissues (layered samples), it is either poorly understood or not used by NIR
94 spectroscopists working with agricultural, food, pharmaceutical and other industrial
95 samples. To our knowledge, polarized NIRS techniques have never been applied to
96 routine or in-line analysis to reduce scattering effects on spectra on turbid media.

97 In this paper, the effectiveness of this multi-scattering correction based on the
98 polarization subtraction is evaluated using a two-component model powder system.
99 The objectives of this paper were to assess the effect of multi-scattering correction (i)
100 on the performances of a calibration model and (ii) on the robustness of the
101 prediction model built from the corrected spectra for predicting the absorber's
102 concentration of powder samples

103 **2. Theoretical Model : Polarization subtraction**

104 Capital bold characters will be used for matrices, e.g. \mathbf{X} ; non bold characters will be
105 used for column vectors, e.g. X
106 Polarization subtraction technique²³⁻²⁵ is based on the polarization-maintaining
107 property of weakly scattered light. When polarized light illuminates a scattering
108 medium, weakly scattered light will emerge in its original polarization state^{21,26-28},
109 while multiple scattered light will emerge with random polarization. In the case of
110 linearly polarized source, the light that is remitted in the same polarization channel as
111 the input illumination is composed by light that has maintained its original
112 polarization state plus a component from the randomly polarized heavily scattered
113 light (cf equation 1). Light that emerges in the orthogonal polarization channel
114 contains only randomly polarized light, approximately equal to the randomly
115 polarized component in the original polarization state (cf equation 2).

$$116 \quad I_{\parallel}(\lambda) = \frac{\Omega}{2\pi} \cdot I_0(\lambda) \cdot S(\lambda) + \frac{\Omega}{2\pi} \cdot I_0(\lambda) \cdot \alpha(\lambda) \cdot M(\lambda) \quad (1)$$

$$117 \quad I_{\perp}(\lambda) = \frac{\Omega}{2\pi} \cdot I_0(\lambda) \cdot \beta(\lambda) \cdot M(\lambda) \quad (2)$$

118

119

120 Where $I_{\parallel}(\lambda)$ and $I_{\perp}(\lambda)$ are the light scattered by the media with parallel and
121 perpendicular polarization respect to the polarization of the illumination light. $I_0(\lambda)$
122 is the intensity of the illumination light. Ω is the collection solid angle, residual
123 term of the integration on the solid angle²³⁻²⁵, of the optical device. $S(\lambda)$ and $M(\lambda)$
124 are the probabilities of light undergoing single and multiple scattering respectively.

125 Since all light undergoes scattering :

126 $S(\lambda) + M(\lambda) = 1$ (3)
127

128 Finally, $\alpha(\lambda)$ and $\beta(\lambda)$ are the multiple light scattered ratio by the media with the
129 parallel and perpendicular polarization respect to the polarization of the illumination
130 light. The sum $\alpha(\lambda)$ and $\beta(\lambda)$ must be one:

131
132 $\alpha(\lambda) + \beta(\lambda) = 1$ (4)
133

134 By subtracting (1) to (2), the intensity of light undergoing single scattering ($I_{ss}(\lambda)$) is
135 equal to :

136 $I_{ss}(\lambda) = I_{\parallel}(\lambda) - I_{\perp}(\lambda) = \frac{\Omega}{2\pi} \cdot I_0(\lambda) \cdot [S(\lambda) + (\alpha(\lambda) - \beta(\lambda)) \cdot M(\lambda)]$ (5)

137 In conclusion, the part of the single-scattering effect in the signal is preserved and the
138 multi-scattering effect is highly reduced.

139

140 **3. Materials and Methods**

141 **3.1 Instrumentation**

142 In the experimental setup (figure 1), a halogen light source (150 W, Leica Cls) was
143 coupled with a 940 μm core diameter optical fiber (numerical aperture (N.A) 0.22,
144 Sedi & ATI). The light delivered by the fiber was collimated by an aspheric lens
145 (F220SMA-B - Thorlabs). The incident beam was a 2 cm diameter circular spot with
146 1° divergence. The incident and reflected beam were polarized through two broad-
147 band (400 nm - 800nm) polarizers (NT52-557, Edmunds Optics). Incident light was

148 linearly polarized and reflected light was collected in a narrow cone (1°). The output
149 from the analyzer was coupled inside an optical fiber (N.A 0.22, Sedi & ATI) by an
150 aspheric lens (F220SMA-B - Thorlabs). This fiber was connected to a spectrometer
151 (MMS1, Zeiss) featuring a detection range of 400 nm - 800 nm, with 3 nm
152 resolution. A constant angle of 70° was maintained between the excitation and
153 collection arms. This angle was chosen to optimize intensity of the reflected beam
154 and to avoid specular reflection.

155 **3.2 Experimental design and sample preparation**

156 Powdered samples mixing sand (Fontainebleau sand VWR International) and
157 coloring dyes (brilliant blue FCF-E133 and chlorophyllin E141, purchased from
158 Colorey, respectively named E133 and E141 in the text) were prepared. Two sand
159 particle size classes were used: S_1 with a diameter less than 250 μm and S_2 with a
160 diameter greater than 250 μm . Sand play the role of a scattering but non absorbing
161 matrix. One or both of the coloring dyes have been added at different densities to the
162 sand, playing the role of absorbing substance in the mixture. Note that absorbers in
163 powdered form also have scattering properties. Particle sizes of the coloring powders
164 were less than 50 μm , with E133 being about three times smaller than E141.

165 Overall, 42 samples were prepared for spectral acquisition composing a calibration
166 set and 12 samples were prepared, afterward and with the same procedure to create
167 an independent test set.

168 The range of sample's colorant densities (in g.L-1) is specified in figure 2.

169 Each sample was directly prepared in an airtight plastic container of 100 mL by
170 adding the precisely weighted corresponding amount of colorant in 20 g of sand
171 using an analytical balance (Kern 770). The maximum dye volume added was not
172 higher than 4% of the total sand volume. Considering that the pore volume for sand
173 is about 40%, and that the dye particles are between 5 to 10 times smaller than the
174 sand, one can make the assumption that the dye would fill the interstices between the
175 sand particles and therefore not increase the initial volume of sand. The density of
176 the colorant d_{dye} in a sample was obtained from :

$$177 \quad d_{dye} = \frac{m_{dye} \cdot d_{sand}}{m_{sand}} \quad (6)$$

178 With m_{dye} the added mass of dye, d_{sand} the density of sand (which differ for S_1 and
179 S_2) and m_{sand} the mass of sand (here 20 g). The colorant density ranged from [0 - 18
180 $g \cdot L^{-1}$].

181 To ensure homogeneity of the mixture, the sample was agitated after preparation and
182 again just before it was carefully transferred in an adapted 5 cm of diameter cup to
183 get an even and horizontal surface.

184 **3.3 Spectral acquisition**

185 For each sample, light was measured with the polarized spectrometer with parallel
186 and perpendicular respect to the polarization of the illumination light. Dark current
187 (I_b) was recorded from all measured spectra and subtracted.

188 A broadband dielectric mirror (BB3-E02, Thorlabs) was used as a reference (I_0) to
189 standardize spectra from non-uniformities of all components of the instrumentation
190 (light source, fibers, lens, polarizer and spectrometer).

191 From these measurements and the equation (5), a raw reflectance (R_w) and a
192 corrected reflectance (R_c), for each sample, were calculated :

$$193 \quad R_w(\lambda) = \frac{(I_{\parallel}(\lambda) - I_{b\parallel}(\lambda)) + (I_{\perp}(\lambda) - I_{b\perp}(\lambda))}{(I_{o\parallel}(\lambda) - I_{b\parallel}(\lambda))} \quad (7)$$

$$194 \quad R_c(\lambda) = \frac{(I_{\parallel}(\lambda) - I_{b\parallel}(\lambda)) - (I_{\perp}(\lambda) - I_{b\perp}(\lambda))}{(I_{o\parallel}(\lambda) - I_{b\parallel}(\lambda))} = \frac{I_{ss}(\lambda) - (I_{b\parallel}(\lambda) + I_{b\perp}(\lambda))}{(I_{o\parallel}(\lambda) - I_{b\parallel}(\lambda))} \quad (8)$$

195

196 With $I_{\parallel}(\lambda)$ and $I_{\perp}(\lambda)$ the intensities of light scatters by the media with parallel and
197 perpendicular polarization respect to the polarization of the illumination light and
198 $I_{o\parallel}(\lambda)$ and $I_{o\perp}(\lambda)$, the intensities of light reflected by the standard mirror with
199 parallel polarization respect to the polarization of the illumination light (as the
200 perpendicular component emerging from the mirror is zero). $I_{b\parallel}(\lambda)$ and $I_{b\perp}(\lambda)$ are the
201 dark current intensities recorded for each measurement.

202

203 **3.4 Multivariate analysis**

204 All computations and multivariate data analysis were performed with Matlab
205 software v. R2012b (The Mathworks Inc., Natick, MA, USA).

206

207 **3.5 Linear unmixing**

208 The first step of Classical Least Square (CLS)²⁹ was used to extract the reflectance
209 *pure spectra* of the absorbers from the reflectance spectra of the mixtures. Also
210 called **K**-matrix, this linear unmixing assumes that a spectrum is a linear combination
211 of the pure component's spectra. The whole calibration set **R** (42 samples mixing the

212 absorbers at different concentrations) and \mathbf{C} , the matrix of sample components
213 concentrations, were used to compute the linear least square estimated $\hat{\mathbf{K}}$ -matrix of
214 the two pure active components (E133 and E141) composing \mathbf{K} knowing that:

$$215 \quad \hat{\mathbf{K}} = \mathbf{RC}^T(\mathbf{CC}^T)^{-1} \quad (9)$$

216

217 Both the raw reflectance spectra (\mathbf{R}_w) and the corrected reflectance spectra (\mathbf{R}_c) were
218 used to compute respectively $\widehat{\mathbf{K}}_w$ and $\widehat{\mathbf{K}}_c$ containing the demixed pure spectra of
219 E133 and E141.

220

221 **3.6 Calibration**

222 Partial Least Square (PLS)³⁰ algorithm was used to model the chemical composition
223 of the powder mixture using \mathbf{R}_w and \mathbf{R}_c . A general PLS model was built using the
224 whole calibration set (42 samples) to predict the samples of the independent test set
225 (12 samples). Secondly, to assess the robustness of the prediction models regarding
226 sand particle size, a PLS model was built with the samples set S_2 and tested on the
227 independent test set S_1 (figure 2). The number of latent variables was determined by
228 comparing performances by leave-one-out cross-validation³¹. Performances (R^2 ,
229 Standard Error of prediction corrected from the bias (SEPC)) and number of latent
230 variables of the different prediction models built with uncorrected and corrected
231 signals were compared when applied on the test set.

232

233 **4. Results and discussion**

234 **4.1 Spectra analysis**

235 *Bulk colorant*

236 The raw spectra $R_w(\lambda)$ and the corrected spectra $R_c(\lambda)$ of the pure powder colorant
237 E133 and E141 are represented in figure 3 (a) and (b).

238 First comment is that the polarization subtraction reduces the global reflectance
239 intensity of the measured signal (by 10 times). It is an expected result as only a small
240 part of the signal is selected: the single-scattered one. Despite this reflectance loss,
241 the corrected spectrum is not noisy and contains information about the sample.

242 Between 400 nm and 700 nm, the raw spectrum and the corrected spectrum have
243 similar shapes. For example, the spectroscopic signature of the colorant E133
244 appears to be purple (as seen in powdered form), mixing a reflectance peak at 450
245 nm (Blue) and at 650 nm (Red). However, these peaks are more marked in the
246 corrected spectra. For the raw spectrum, crushing peaks can be explained by a strong
247 increase in reflectance after 750 nm. This sharp increase in reflectance, in a spectral
248 range where the colorant does not absorb, is due to the multi-scattering. However,
249 this effect seems to be less important for E141 than for E133. In the corrected
250 spectra, this effect is strongly reduced (figure 3 (b)).

251

252 *Sand and dye mixtures*

253 Figures 3 (c) and (d) show respectively the raw spectra and the corrected spectra of
254 sand S_1 mixing coloring powder E133 at different densities. When mixed with the
255 colorant, sand is responsible for high multi-scattering as it is not absorbing the light.

256 This physical phenomenon comes on top of the chemical information contained in
257 the spectra and masks the spectral features of the absorber. The shape of the raw
258 spectra of coloring powder E133 (figure 3 (a)) and the shapes of the raw reflectance
259 of the sand-dye mixture (figure 3 (c)) are completely different. The multiplicative
260 effect due to scattering is not wavelength dependent and depends on the number of
261 scatterers in the sample. Therefore, the more dye, the more scatterers and,
262 consequently the higher the reflectance. The order of the raw spectra is consistent
263 with the dye's concentration, but for a physical reason.

264 By applying the correction, the spectral features of the colorant are enhanced as it
265 can be seen on figure 3 (d). The signature of the corrected spectra is similar to the
266 spectral signal of the colorant E133 in powdered form (figure 3 (b)). The reflectance
267 peaks at 450 nm and at 650 nm clearly appear, but more important, because linked to
268 the dye's concentration, the wavelength regions where absorbance occurs (400- 430
269 nm and 500 – 650 nm) are now visible. In these regions, the spectrum ordering is
270 consistent with the dye concentration, contrary to the other areas where low
271 absorbance occurs and reveals more complex reflectance patterns.

272 Figures 3 (e) and (f) show respectively the raw spectra and the corrected spectra of
273 sand S_1 mixing this time coloring powder E141 at different densities. By comparing
274 the raw reflectance intensities of the sand - E133 mixtures (figure 3 (c)) and sand -
275 E141 mixtures (figure 3 (e)), containing the same ranges of dye's densities, it appears
276 that the level of R_w is two-times higher for E133 than for E141. As stated in section
277 'Experimental design and sample preparation', the particle size of E141 is, at least,

278 three-times larger than E133. This difference in particle diameter has a direct impact
279 on the elastic scattering phenomenon occurring during light-matter interaction. First,
280 for the same density of dye, small particles scatter more than larger particles²¹.
281 Secondly, the scattering angle differs between small and large particles: the larger the
282 diameter, the smaller the scattering angle. Combining these two properties, the
283 overall reflectance intensity will be higher for smaller particles, which is the case of
284 raw reflectance of the Sand-E133 mixture.

285 For the corrected spectra R_c , there is no significant difference in the intensity level
286 between sand - E133 mixtures (figure 3 (d)) and sand - E141 mixtures (figure 3 (f)).
287 This is coherent with the fact that the method corrects the spectra from multi-
288 scattering, which is mainly due to the sand particles but also, and in a significant
289 manner, to the powdered colorant.

290

291 **4.2 Extraction of the absorber's pure spectra**

292 Figure (4) presents the demixed pure spectra \widehat{K} extracted from the reflectance spectra
293 (R_w and R_c). The linear unmixing (figure 4 (a)) applied to raw spectra (R_w) provided
294 estimated pure spectra ($\widehat{K}_{w_E133}(\lambda)$ and $\widehat{K}_{w_E141}(\lambda)$) that are very different from the
295 raw spectra $R_{w_E133}(\lambda)$ and $R_{w_E141}(\lambda)$ measured directly on the powders. The shape
296 of these spectra are not matching and the correlation coefficients, between
297 $\widehat{K}_{w_E133}(\lambda)$ and $R_{w_E133}(\lambda)$ and between $\widehat{K}_{w_E141}(\lambda)$ and $R_{w_E141}(\lambda)$, are respectively
298 equal to 0.59 and 0.74.

299 Linear unmixing assumes that a spectrum is a linear combination of the pure
300 components spectra. From these results, the failure to recover the absorber's pure
301 spectrum from the raw reflectance spectra, support the fact that interactions in the
302 mixture are responsible of non-linearities which are directly responsible of the non-
303 linearities in the relationship between the absorbance and the absorbing power of the
304 sample. It is well known^{20,32} that scattering and absorbance are not independent
305 phenomenon. Scatter increases the mean free path of photons, which increases the
306 chances of being absorbed.

307 On the contrary, the computation of the $\widehat{\mathbf{K}}_c$ -matrix is successful and matches the pure
308 colorant corrected reflectance spectrum. Estimated pure spectra ($\widehat{K}_{c_E133}(\lambda)$ and
309 $\widehat{K}_{c_E141}(\lambda)$) are very close to $R_{c_E133}(\lambda)$ and $R_{c_E141}(\lambda)$ measured directly on the
310 powder and corrected. The correlation coefficients are respectively 0.98 and 0.93. As
311 the method corrects the signal from a physical phenomenon (multi-scattering), one
312 can affirm that the observed interactions in the raw spectra are of optical nature (and
313 not chemical interactions). Regarding these results and this consideration,
314 polarization subtraction correction induces, by reducing the multi-scattering effect, a
315 better linear relationship between the light attenuation and the absorption the dyes.
316 The recovered spectra $\widehat{\mathbf{K}}_c$ are corrected from the physical interactions occurring in
317 the mixture.

318

319 **4.3 Calibration model**

320 *General model*

321 Table 1 shows the quality parameter of the prediction models of the test set
322 absorber's densities comparing the raw and the corrected spectra.

323 First, with the raw spectra $R_w(\lambda)$, the quality of the E133 prediction model is
324 distinctly poorer than the quality of the E141 prediction model. The number of latent
325 variables is much higher for E133 as well as the SEP_c . On the contrary, the figures of
326 merit for the E141 model are good. Again, E133 and E141 behave differently. As
327 stated before, multiscattering effect is more important when E133 is present in the
328 mixture (with or without E141) because of a smaller particle diameter. Hence, non-
329 linearities between absorbance and absorber's concentration are more important and
330 the PLS model is limited in building a performant linear prediction model for E133.

331 When the models are built with the corrected reflectance spectra $R_c(\lambda)$, the figure of
332 merit of the absorber's density prediction models are good and of the same level for
333 E133 and E141. In both cases, the number of latent variables is decreasing. The
334 improvement is less important for E141 but significant for E133. And both models
335 have a lower standard error of prediction compared to the $R_w(\lambda)$ prediction models
336 for the two colorants. In section 'Sand and dye mixture', we stated that the proposed
337 method mainly reduced the effect of multiscattering due to sand particles, enhancing
338 the part of the signal related to chemical absorbance. In $R_w(\lambda)$, while this information
339 is present, it is masked by the multi-scattering and the PLS needs more latent
340 variable to extract this chemically related information to build a model.

341 Here, the different behavior of the two dyes is not obvious anymore. This agrees with
342 the hypothesis that the correction method equalizes the mean free photon path

343 between all the samples, regardless of the particle size and shape of the sample's
344 constituents. For the two absorbers, the number of latent variables, which is an
345 indicator of the complexity of the models, is still high (respectively 5 and 4 for E133
346 et E141 with the corrected spectra $R_c(\lambda)$) for samples mixing only two different
347 absorbers. Theoretically, two PLS components should be sufficient. This agrees with
348 the initial assumption that the polarization subtraction method highly reduces the
349 multi-scattering but does not remove it completely. In addition, these results are
350 consistent with the conclusions of section 'Extraction of the absorber's pure spectra',
351 stating that the correction method restores, in a significant manner, the linear
352 relationship between the spectra and the absorber's density in the powdered samples.
353 To conclude, as the PLS models using the corrected spectra show good prediction
354 capacities, it fulfills the assumption that even if the corrected signal intensity is
355 highly reduced, the remaining information is of better quality in terms of signal
356 sensitivity.

357

358 *Robustness assessment*

359 Table 2 presents the results of the calibration model built with samples of one
360 particle size ($S_2 > 250 \mu\text{m}$) and tested on samples with another particle size ($S_1 < 250$
361 μm).

362 First, the models built with $R_w(\lambda)$ show, as previously observed, better predictions
363 for E141 than for E133, but, in overall, lower quality than in table 1. This confirms a
364 different behavior of E133 and E141, but also that the sand particle size has an effect

365 in the quality of the predictions. A change in the physical structure of the samples
366 usually leads to low prediction performances because of the scattering impact on the
367 signal.

368 When built with the corrected spectra, again, the prediction of E133 highly improves,
369 while the gain is less significant for E141, which is also consistent with the previous
370 conclusions. But overall, the predictions are good, confirming that the corrected
371 spectra, composed by the single scattered part of the total reflectance signal, becomes
372 less dependent to physical changes in the sample.

373 The polarization subtraction method selects by optical means only, part of
374 information related to the powdered absorbers concentration, while discarding the
375 unwanted effect of multi-scattering on the signal. The measured signal becomes less
376 dependent of the particle size changes of the samples and therefore improves both
377 quality and robustness of the prediction models.

378

379 **5. Conclusion**

380 This study demonstrates the effectiveness of the polarized light subtraction method,
381 applied to a two component model powder system, which improves the performance
382 of multivariate calibration models.

383 By selecting only the light which has conserved the initial polarization and therefore
384 being less impacted by scattering events, the collected photons all have the same
385 mean free pathlength. As a consequence, the linear relationship between the spectra
386 and the absorbers concentration is restored.

387 Then it is possible, by using Linear Unmixing to recover the pure spectra of the dyes.
388 These recovered spectra are corrected from the physical interactions occurring in the
389 mixture.

390 When the corrected spectra are used to build the PLS models, all the general quality
391 parameters and the parsimony improve, especially for highly scattering media.
392 Although the overall signal intensity is reduced after optical correction, the
393 remaining information in the corrected signal is sufficient and of better quality to
394 build a good prediction model, thus meaning that the signal sensitivity increases.
395 Inevitably there is a trade-off between making more accurate measurements and a
396 reduction of SNR.

397 After the polarization correction, the measured signal becomes less dependent to
398 physical changes (particle sizes) which also improve the robustness of the prediction
399 models.

400 This "plug and play" optical method offers the potential to be easy to implement to a
401 commercial spectrophotometry system and does not significantly increase the
402 measurement time.

403

404 **References and links**

405 (1) Hart, J.; Norris, K.; Golumbic, C. *Cereal Chem* 1962, 39, 94–99.

406 (2) Massie, D. R.; Norris, K. H. *Transactions of the ASABE* 1965, 8 (4), 598–600.

407 (3) Shi, Z.; Anderson, C. A. *Journal of Pharmaceutical Sciences* 2010, 99, 4766–
408 4783.

- 409 (4) Chauchard, F.; Roger, J.; Bellon-Maurel, V.; Abrahamsson, C.; Andersson-
410 Engels, S.; Svanberg, S. *Applied Spectroscopy* 2005, 59, 1229–1235.
- 411 (5) Jin, J.-W.; Chen, Z.-P.; Li, L.-M.; Steponavicius, R.; Thennadil, S. N.; Yang, J.;
412 Yu, R.-Q. *Analytical Chemistry* 2012, 84, 320–326.
- 413 (6) Rinnan, A.; Berg, F. v. d.; Engelsen, S. B. *TrAC Trends in Analytical Chemistry*
414 2009, 28, 1201–1222.
- 415 (7) Martens, H.; Stark, E. *Journal of Pharmaceutical and Biomedical Analysis* 1991,
416 9, 625–635.
- 417 (8) Barnes, R. J.; Dhanoa, M. S.; Lister, S. J. *Appl. Spectrosc.* 1989, 43, 772–777.
- 418 (9) Geladi, P.; MacDougall, D.; Martens, H. *Appl. Spectrosc.* 1985, 39, 491–500.
- 419 (10) Chen, Z.-P.; Morris, J.; Martin, E. *Analytical Chemistry* 2006, 78, 7674–7681.
- 420 (11) Martens, H.; Nielsen, J. P.; Engelsen, S. B. *Analytical Chemistry* 2003, 75, 394–
421 404.
- 422 (12) Steponavicius, R.; Thennadil, S. N. *Analytical Chemistry* 2011, 83, 1931–1937.
- 423 (13) Prahl, S. *Optical-thermal response of laser-irradiated tissue, A* 1995, 122–127.
- 424 (14) Steponavicius, R.; Thennadil, S. N. *Analytical Chemistry* 2009, 81, 7713–7723.
- 425 (15) Farrell, T.; Patterson, M.; Wilson, B. *Medical Physics* 1992, 19, 879–888.
- 426 (16) Abrahamsson, C.; Johansson, J.; Andersson-Engels, S.; Svanberg, S.; Folestad,
427 S. *Analytical Chemistry* 2005, 77, 1055–1059.
- 428 (17) Torrance, S.; Sun, Z.; Sevick-Muraca, E. *Journal of Pharmaceutical Sciences*
429 2004, 93, 1879–1889.

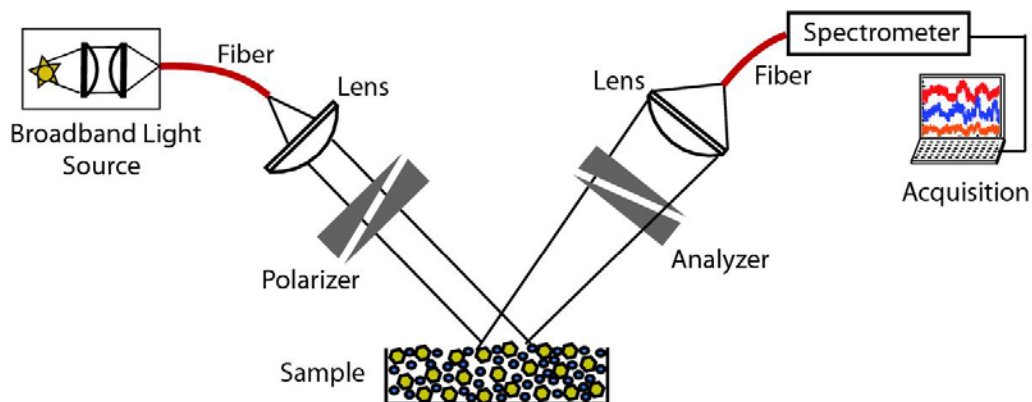
- 430 (18) Swartling, J.; Dam, J.; Andersson-Engels, S. *Applied Optics* 2003, 42, 4612–
431 4620.
- 432 (19) Hebden, J. C.; Arridge, S. R.; Delpy, D. T. *Physics in Medicine and Biology*
433 1997, 42, 825.
- 434 (20) Lu, B.; Morgan, S. P.; Crowe, J. A.; Stockford, I. M. *Appl. Spectrosc.* 2006, 60,
435 1157–1166.
- 436 (21) Backman, V.; Gurjar, R.; Badizadegan, K.; Itzkan, I.; Dasari, R. R.; Perelman,
437 L. T.; Feld, M. *Selected Topics in Quantum Electronics, IEEE Journal of* 1999, 5,
438 1019–1026.
- 439 (22) Abrahamsson, C.; Löwgren, A.; Strömdahl, B.; Svensson, T.; Andersson-
440 Engels, S.; Johansson, J.; Folestad, S. *Appl. Spectrosc.* 2005, 59, 1381–1387.
- 441 (23) Schmitt, J.; Gandjbakhche, A.; Bonner, R. *Appl. Opt.* 1992, 31, 6535–6546.
- 442 (24) Morgan, S.; Ridgway, M. *Optics Express* 2000, 7, 395–402.
- 443 (25) Demos, S. G.; Alfano, R. R. *Appl. Opt.* 1997, 36, 150–155.
- 444 (26) Demos, S. G.; Alfano, R. R. *Opt. Lett.* 1996, 21, 161–163.
- 445 (27) Sokolov, K.; Drezek, R.; Gossage, K.; Richards-Kortum, R. *Opt. Express*
446 1999, 5, 302–317.
- 447 (28) Yoo, K.; Alfano, R. *Physics Letters A* 1989, 142, 531–536.
- 448 (29) Geladi, P. *Spectrochimica Acta Part B: Atomic Spectroscopy* 2003, 58, 767 –
449 782.
- 450 (30) Wold, S.; Sjöström, M.; Eriksson, L. *Chemometrics and Intelligent Laboratory*
451 *Systems* 2001, 58, 109–130.

452 (31) Wold, S. *Technometrics* 1978, 20, 397–405.

453 (32) Stockford, I. M.; Lu, B.; Crowe, J. A.; Morgan, S. P.; Morris, D. E. *Appl.*

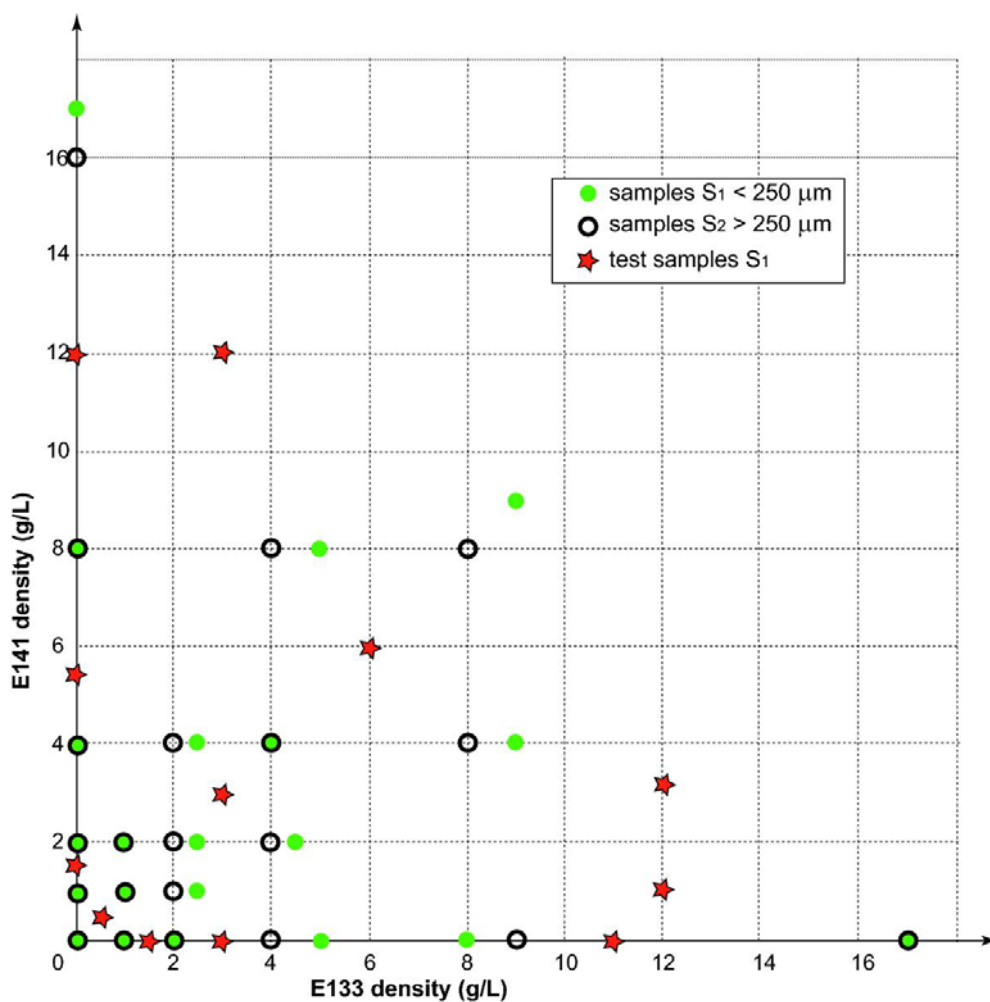
454 *Spectroscopy* 2007, 61, 1379–1389.

455



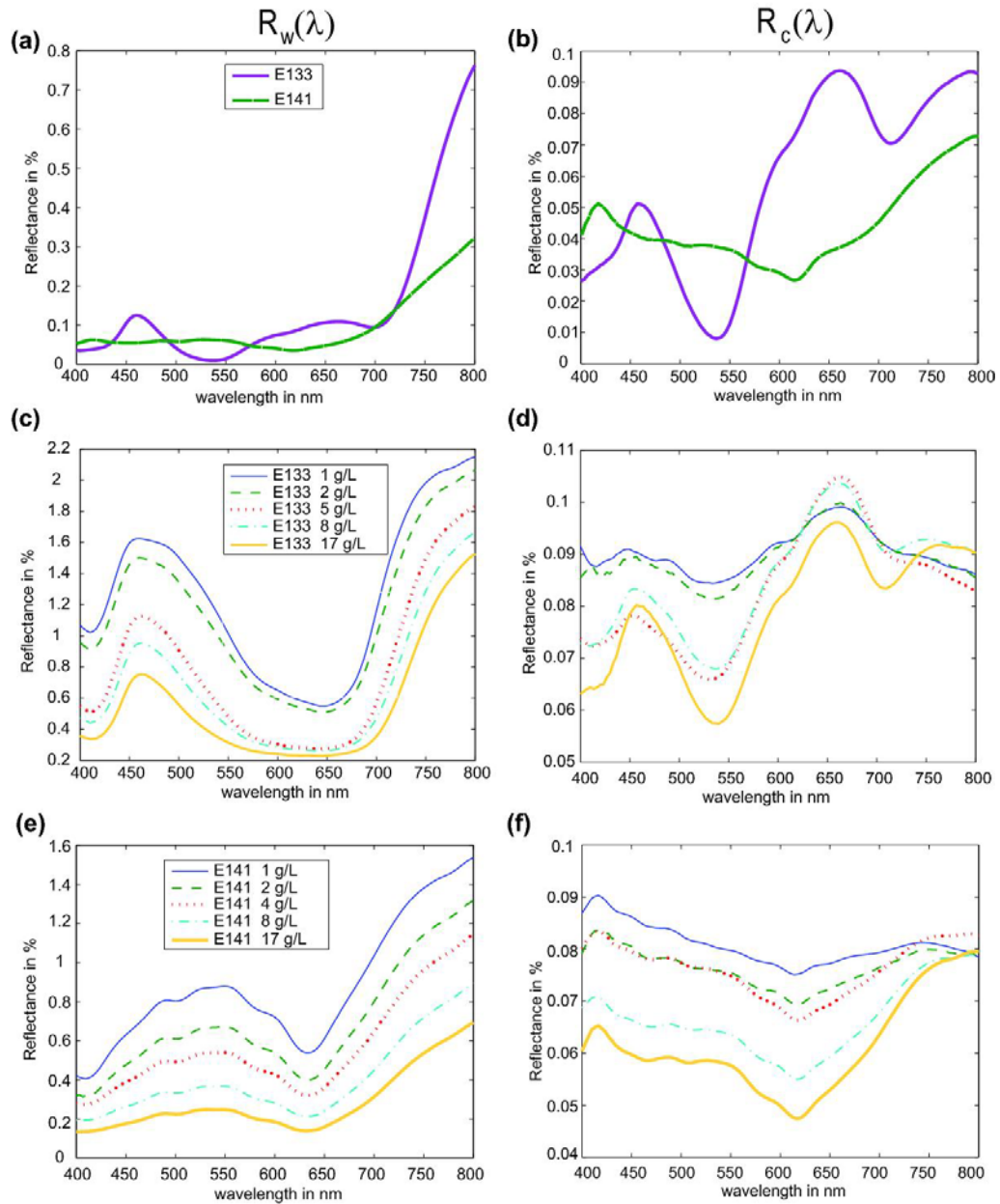
456
457
458
459

Figure 1: Schematic diagram of polarized light spectroscopy system.



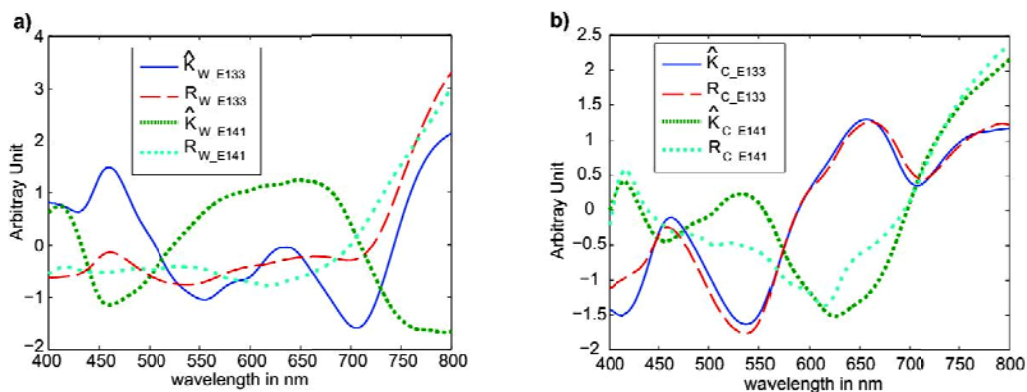
460
461
462
463

Figure 2: Experimental design presenting the dye densities in g.L^{-1} of 42 samples for the calibration set and 12 samples for the independent test set.



464
465
466
467
468
469
470
471
472
473

Figure 3: (a) Raw reflectance spectra of coloring powder E133 and E141. (b) Corrected reflectance spectra of coloring powder E133 and E141 (c) Raw reflectance spectra of sand S_1 + coloring powder E133 mixed at different densities. (d) Corrected reflectance spectra of sand S_1 + coloring powders E133 mixed at different densities.(e) Raw reflectance spectra of sand S_1 + coloring powder E141 mixed at different densities.(f) Corrected reflectance spectra of sand S_1 + coloring powders E141 mixed at different densities.



475
476

480 Figure 4 : Comparison of the raw and the corrected spectra acquired on the two
 481 coloring powders ($R_{w_E133}(\lambda)$, $R_{w_E141}(\lambda)$ and $R_{c_E133}(\lambda)$, $R_{c_E141}(\lambda)$) with the
 482 demixed pure spectrum ($K_{w_E133}(\lambda)$, $K_{w_E141}(\lambda)$ and $K_{c_E133}(\lambda)$, $K_{c_E141}(\lambda)$) extracted
 483 respectively from (R_w) and (R_c) with Linear unmixing.

481
482
483
484
485

Table 1: Figure of merit of the calibration models

	Cal set (n)	Test set (n)	Spectra	Predicted absorber	PLS factors	R ²	SEPC (% g.L ⁻¹)
General model	S ₁ +S ₂ (42)	S ₁ (12)	R_w	E133	8	0.75	2.52
				E141	5	0.91	1.3
			R_c	E133	5	0.91	1.41
				E141	4	0.93	1.12

486
487
488

490 Table 2: Figure of merit of the calibration model built with samples of one particle
 491 size (S₂) and tested on samples with another particle size (S₁)

491

	Cal set (n)	Test set (n)	Spectra	Predicted absorber	PLS factors	R ²	SEPC (% g.L ⁻¹)
General model	S ₂ (21)	S ₁ (12)	R_w	E133	5	0.69	3.33
				E141	5	0.83	1.79
			R_c	E133	4	0.91	1.72
				E141	3	0.93	1.73

492