



**HAL**  
open science

# A simple correction method of inner filter effects affecting FEEM and its application to the PARAFAC decomposition

Xavier Luciani, Stéphane Mounier, Roland Redon, André Bois

► **To cite this version:**

Xavier Luciani, Stéphane Mounier, Roland Redon, André Bois. A simple correction method of inner filter effects affecting FEEM and its application to the PARAFAC decomposition. *Chemometrics and Intelligent Laboratory Systems*, 2009, 96 (2), pp.227-238. 10.1016/j.chemolab.2009.02.008 . hal-00807769

**HAL Id: hal-00807769**

**<https://hal.science/hal-00807769>**

Submitted on 4 Apr 2013

**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.

# A simple correction method of inner filter effects affecting FEEM and its application to the PARAFAC decomposition

X. Luciani<sup>a,b,\*</sup>, S. Mounier<sup>b</sup>, R. Redon<sup>b</sup>, A. Bois<sup>b</sup>

<sup>a</sup>*Laboratoire I3S, UMR6070, UNSA CNRS, 2000, route des Lucioles, BP 121, 06903 Sophia Antipolis Cedex - France*

<sup>b</sup>*Laboratoire PROTEE, USTV, BP 20132, 83957 La Garde Cedex, France*

---

## Abstract

In this paper we introduce a new inner filters correction method for standard fluorometer. The Controlled Dilution Approach (CDA) deals with highly absorbing solutions using the Fluorescent Excitation-Emission Matrix of a controlled weak dilution. Along with the non linear FEEM of the original solution, these two informations allow to estimate the linearized FEEM. The method relies on inner filter effects modelization. Beyond its numerical simplicity, the main interest is that CDA only requires fluorescence measurements. The method was validated using a set of known mixtures and a set of dissolved organic matter samples. In addition we show that the corrected FEEM can be used efficiently for advanced multilinear analysis. Therefore CDA is presented here as a relevant pretreatment to the PARAFAC decomposition of highly absorbing mixtures.

*Key words:* EEM, inner filter effects, PARAFAC, 3D fluorescence

---

---

\* Corresponding autor.

*Email address:* lucianix@gmail.com (X. Luciani).

# 1 Introduction

## 1.1 Non linearities in fluorescence spectroscopy

Fluorescent molecular components (fluorophores) can be easily distinguished by their spectroscopic properties and more particularly by their fluorescence spectra [1]. Recent fluorimeters provide successive measurements of the fluorescence intensity emitted by a solution of one or several fluorophores. By scanning excitation and emission wavelength domains, Fluorescent Excitation-Emission Matrices (FEEM) gather a lot of informations about the solution. These spectra are now widely used in various scientific domains such as medicine [2], analytical chemistry [3] or environmental sciences [4,5]. Ideally, considering the FEEM ( $I_{3D}$ ) of a single fluorophore, its norm is proportional to the fluorophore concentration in the solution and its pattern is given by the outer product between the excitation spectrum and the emission spectrum of the fluorophore. This is the classical linear model of fluorescence. However it is well known that its pertinence decreases with the concentration [6]. Actually, non linear deviations mean that the gradual absorption by the solution of both exciting and fluorescent lights cannot be neglected. These effects are known as inner filter effects and affect both  $I_{3D}$  norm and  $I_{3D}$  pattern. Therefore, in presence of inner filter effects, one cannot deduce any correct information about the solution directly from  $I_{3D}$ .

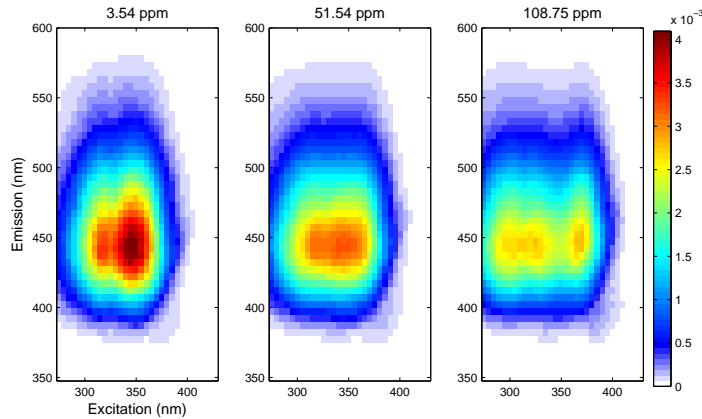


Fig. 1. Evolution of the quinine sulphate 3D spectrum for three different concentrations: 3.54 ppm (left spectrum), 51.54 ppm (middle spectrum), 108.75 ppm (right spectrum)

Example of inner filter effects is given on figure 1. This example clearly shows that the FEEM pattern can be severely affected even in the simple case of a single fluorophore solution.

In other respects, considering several solutions of the same diluted fluorophore

26 measured in different conditions, many other factors such as diffusion, tem-  
 27 perature variations, pH variations, fluorescence quenching or ionic strength  
 28 can affect the FEEM linearity [1]. In this work, we only focus on inner filter  
 29 effects correction.

30

## 31 1.2 Inner filter effects correction

32 Inner filter effects are observed and studied for a long time now [7,8]. Two  
 33 main correction methods are used to prevent these deviations. Since inner  
 34 filter effects can be neglected for weak absorbances *i.e.* weak concentrations, a  
 35 common procedure is to strongly dilute the solution until maximal absorbance  
 36 is inferior than 0.1 [1]. There is an obvious drawback with this dilution method  
 37 as a too strong diluting factor would severely reduce the signal to noise ratio.  
 38 Moreover this procedure must be applied very carefully to avoid contamination  
 39 or physico-chemical changes. Therefore, ensuring the linearity of the data set  
 40 is no easy task. The second approach uses a mathematical model of inner filter  
 41 effects [9–11]. Then one can deduce a correction factor in order to estimate  
 42 element by element, a corrected FEEM ( $I_c$ ) from  $I_{3D}$ . It is assumed that if the  
 43 correction factor is suitable then  $I_c$  will follow the linear model. This approach  
 44 relies on the Beer-Lambert law [1] which gives the elementary variation  $dI$  of  
 45 the light intensity through an elementary optical path  $dl$  at wavelength  $\lambda$ :

$$dI = -I(\lambda)\alpha(\lambda)dl \quad (1)$$

46 where  $\alpha$  is the absorption coefficient of the solution. Then, the integrated  
 47 law describes the light absorption through the entire optical path. If  $I_0$  is the  
 48 intensity of the exciting light, the transmitted intensity outside a cell of length  
 49  $l$  is simply given by the relation:

$$I(\lambda) = I_0(\lambda)e^{-\alpha(\lambda)l} = I_0(\lambda)10^{-A(\lambda)} \quad (2)$$

50 The absorbance spectrum of the solution is then defined by

$$A(\lambda) = \log_{10} \left( \frac{I_0(\lambda)}{I(\lambda)} \right) = \frac{l\alpha}{\log(10)} \quad (3)$$

51 Figure 2 shows a schematic diagram of absorbance measurement.  $A$  is obtained  
 52 by measuring the transmitted intensity through the diluted solution ( $I_T$ ) and  
 53 the transmitted intensity through the solvent ( $I_R$ ) at successive wavelength:

$$A(\lambda) = \log_{10} \left( \frac{I_R(\lambda)}{I_T(\lambda)} \right) \quad (4)$$

54 In right angle fluorescence spectroscopy, classical model of inner filter effects  
 55 is given by equation 5.

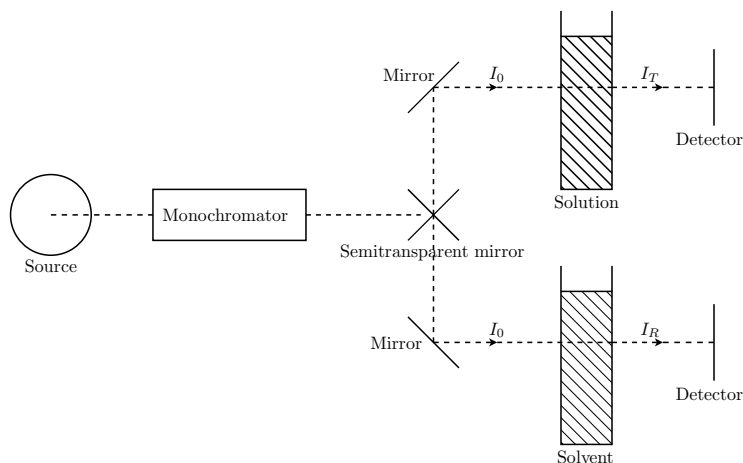


Fig. 2. Absorbance measurement

$$I_{3D}(\lambda_{ex}, \lambda_{em}) = I_c 10^{-\frac{A(\lambda_{ex}) + A(\lambda_{em})}{2}} \quad (5)$$

56 Thereby, one can use the measured absorbance spectrum to compute the cor-  
 57 rection factor and then deduce  $I_c$ . Similar methods were proposed in [12–14].  
 58 In the following, this approach will be identified as the Absorbance Correction  
 59 Approach (ACA). ACA is commonly used in applicative papers dealing with  
 60 fluorescence spectroscopy [15,16]. However absorbance measurement is much  
 61 less sensitive than fluorescence measurement. In addition it requires an other  
 62 experimental device whose characteristics are different, introducing its own  
 63 error in the chain. Finally, the short linear range of absorption measurement  
 64 is another important drawback of ACA. In this work we propose an original  
 65 correction method: the Controlled Dilution Approach (CDA) which combines  
 66 the advantages of both methods. CDA uses the FEEM of a diluted solution  
 67 instead of absorbance measurement in order to estimate  $I_c$ . The crucial point  
 68 is that the dilution factor can be chosen small enough to avoid the drawback  
 69 of the dilution approach. Indeed, the linearity of this second FEEM is not re-  
 70 quired. Consequently, CDA keeps the main advantage of ACA which is a very  
 71 simple numerical correction, but it only requires fluorescence spectra. Anal-  
 72 yzed solutions are generally mixtures of several fluorophores. Therefore, many  
 73 applications involve a separation step to recover the underlying individual  
 74 spectra and concentration profiles of each fluorophore. Number of chemomet-  
 75 ric methods were proposed in the literature in order to perform multilinear  
 76 decompositions of FEEM [17–20]. Based on original works of Harshman [21],  
 77 PARAllel FACtor analysis (PARAFAC) was introduced in this context by Bro  
 78 [22]. During the last decade PARAFAC has proved to be the most relevant  
 79 approach. For instance, in environmental sciences, it is currently the reference  
 80 tool to characterize and trace Dissolved Organic Matter (DOM) [23–25]. In  
 81 return, it do not take into account inner filter effects [26,27]. Consequently  
 82 there is an irreversible loss of performance when dealing with highly absorb-  
 83 ing mixtures.

84 Like other inner filter effects correction methods, CDA is independent of this  
85 separation step. However, we take into consideration that a large part of FEEM  
86 applications, uses this kind of decomposition. As a consequence, in order to  
87 ensure the reliability of CDA, we also present in this paper its performance as  
88 a PARAFAC pretreatment of highly absorbing mixtures.

89

### 90 1.3 Paper organization

91 CDA is detailed on section 2 of this paper. First the modelization of inner  
92 filter effects is given in section 2.1 then CDA is described on section 2.2.  
93 Lastly, practical aspects of CDA are presented in section 2.3 notably in the  
94 case of FEEM sets analysis. The PARAFAC application to the CDA corrected  
95 FEEM is shortly describe.

96 In this work, CDA correction is experimentally tested on two very different  
97 sets of mixtures. The first set is composed of standard laboratory mixtures  
98 of fluorescein and quinine sulphate. Consequently, thist first data set is used  
99 to strictly validate CDA and compare with classical ACA. On the other side,  
100 the second data set is constituted by unknown samples of DOM catchments  
101 and gives an example of how the method can help in a realistic case. Section  
102 3 describes the experimental part of these tests. Results obtained on both  
103 data sets before and after the PARAFAC decomposition are presented and  
104 discussed in section 4.

## 105 2 Theory

### 106 2.1 Modelization of inner filter effects

107 Like ACA, CDA relies on equation 5. Few authors give detailed mathematical  
108 justifications of this model, particularly in the most general case of 3D spectra  
109 of fluorophore mixtures. In this subsection a rigorous interpretation of equa-  
110 tion 5 is proposed.

111 We consider here a mixture of  $N$  fluorophores. For each fluorophore  $n$ , we  
112 note  $c_n$  its concentration in the solution,  $\varepsilon_n(\lambda_{ex})$  its molar extinction coeffi-  
113 cient at the excitation wavelength  $\lambda_{ex}$ ,  $\Phi_n$  its the quantum yield of fluorophore  
114  $n$ ,  $\gamma_n(\lambda_{em})$  its emission probability at wavelength  $\lambda_{em}$  and  $\alpha_n(\lambda_{ex})$  its absorp-  
115 tion coefficient which is equal to the product of  $c_n$  by  $\varepsilon_n(\lambda_{ex})$ . We assume  
116 that the absorption and emission spectra of fluorophore  $n$  are normalized val-  
117 ues of respectively  $\varepsilon_n(\lambda_{ex})$  and  $\gamma_n(\lambda_{em})$ . In the linear approximation, every  
118 fluorescing particles are treated equally as if the whole sample cell was an ele-

119 mentary point. In order to improve this model, one should takes into account  
 120 the particular geometry of the problem.

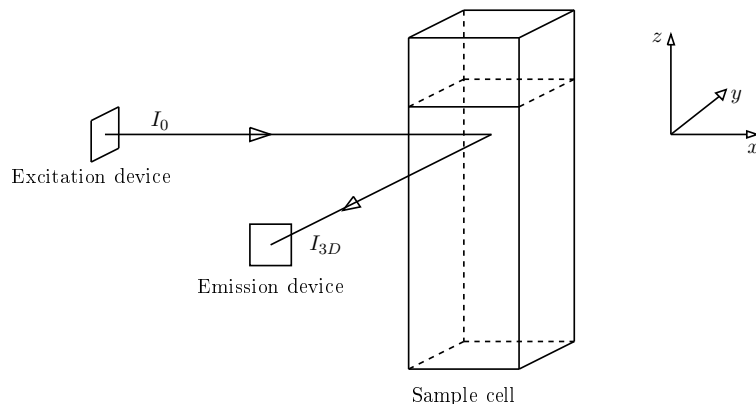


Fig. 3. Schematic diagram of right angle fluorescence measurement

121 Figure 3 recalls basically the experimental device of right angle standard flu-  
 122 orimeters. The excitation light ( $I_0(\lambda_{ex})$ ) is absorbed through the sample cell  
 123 (length  $l$ ) by the fluororophores, inducing the fluorescent light. Finally, a frac-  
 124 tion ( $I_{3D}(\lambda_{em})$ ) of the emitted signal is collected perpendicularly to the excit-  
 125 ing beam.  $\lambda_{ex}$  and  $\lambda_{em}$  scannings allow to measure the FEEM.

126 In this study, several approximations were made. First of all, we took into  
 127 consideration the symmetry of the problem, therefore the influence of the  $z$   
 128 spatial dimension was neglected.

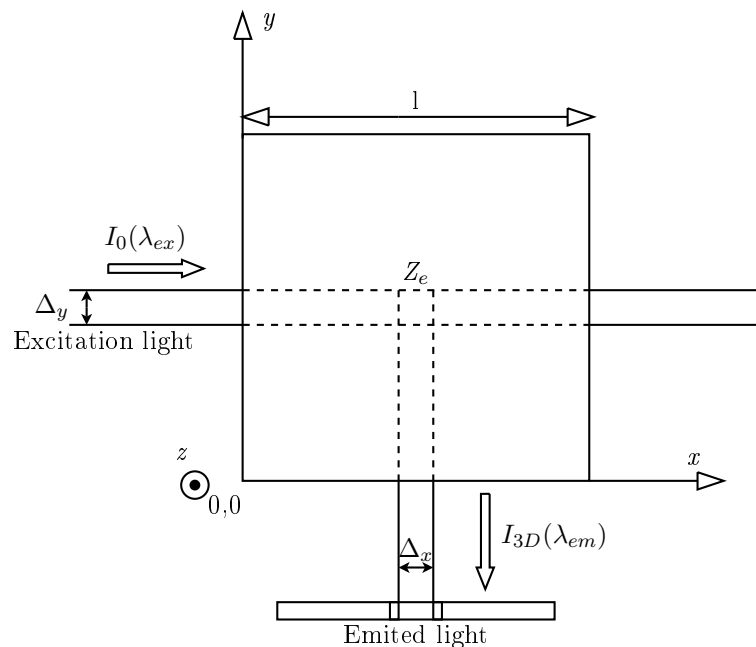


Fig. 4. Scheme of the sample cell, view from above

129 Secondly, only two main optical paths were considered. They represent the  
 130 excitation beam and emission beam in figure 4 scheme. This means that the  
 131 fraction of the exciting light which do not reach the "influence zone"  $Z_e$  was  
 132 neglected as well as the fluorescence light issued from the region outside  $Z_e$ .  
 133 Then each elementary segment of the "excited face" of  $Z_e$  was supposed to  
 134 receive the same energy from the rectilinear exciting beam. In the same way,  
 135 we assumed that each elementary segment of the "emission face" of  $Z_e$  provides  
 136 the same energy to the detector. Furthermore, diffusion and re-emission effects  
 137 were also neglected. Actually we only consider the elementary optical paths  
 138 represented in figure 5.

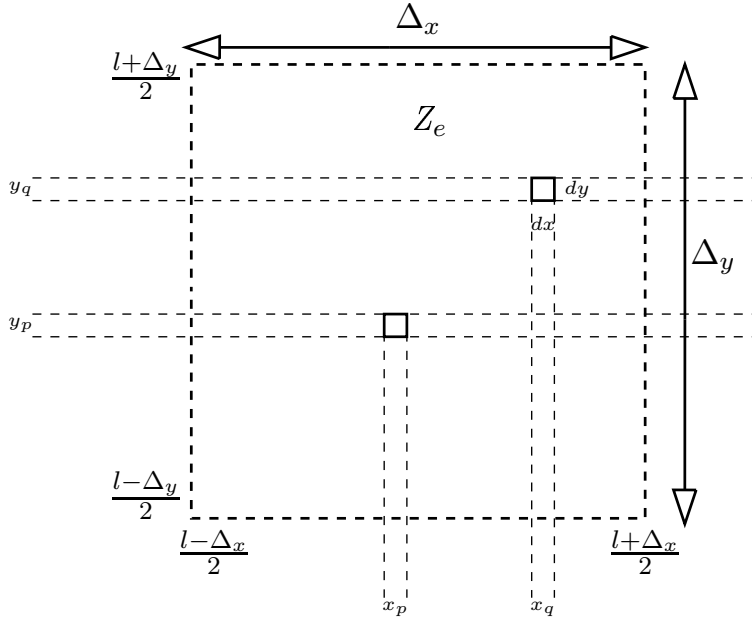


Fig. 5. Elementary cutting of the "influence zone" in the directions of exciting and emitted lights

139 The integrated Beer-Lambert law describes the light absorption through the  
 140 optical path. If  $I_0$  is the light intensity at the point  $x_0$  of the dilute solution,  
 141 the intensity in  $x$  is simply given by the relation:  $I = I_0 e^{-\alpha(\lambda_{ex})(x-x_0)}$  where  $\alpha$   
 142 is the mixture total absorption coefficient:  $\alpha = \sum_n \alpha_n$ . The influence zone was  
 143 divided into horizontal and vertical elementary strips of respective dimension  
 144  $dy \times l$  and  $l \times dx$ . Then each horizontal strip receives an equal elementary  
 145 fraction of the exciting light:  $\frac{dy I_0}{\Delta y}$  and the Beer-Lambert law quantifies the  
 146 intensity transmitted to  $x$ . A fraction  $\alpha_n dx$  is absorbed by fluorophore  $n$ , and  
 147 the total intensity absorbed by fluorophore  $n$  in the "influence zone" ( $I_{An}$ ) is  
 148 given by:

$$I_{An}(\lambda_{ex}) = \frac{dy I_0}{\Delta y} \alpha_n(\lambda_{ex}) \int_{\frac{l - \Delta x}{2}}^{\frac{l + \Delta x}{2}} e^{-\alpha(\lambda_{ex})x} dx \quad (6)$$



$$I_{An}(\lambda_{ex}) = 2 \frac{I_0}{\Delta_y \alpha(\lambda_{ex})} \alpha_n(\lambda_{ex}) e^{-\frac{\alpha(\lambda_{ex})l}{2}} \sinh\left(\frac{\alpha(\lambda_{ex})\Delta_x}{2}\right) dy \quad (7)$$

149 The fluorescence signal emitted by the  $dy$  strip at wavelength  $\lambda_{em}$  is equal  
 150 to  $\sum_n \Phi_n \gamma_n(\lambda_{em}) I_{An}(\lambda_{ex})$  and the Beer-Lambert law integrated on all the ele-  
 151 mentary horizontal strips gives the ratio of the fluorescence signal transmitted  
 152 outside the sample cell in the  $y$  direction,  $I_{3D}(\lambda_{ex}, \lambda_{em})$ .

$$I_{3D}(\lambda_{ex}, \lambda_{em}) = \int_{\frac{l-\Delta_y}{2}}^{\frac{l+\Delta_y}{2}} \sum_n \Phi_n \gamma_n(\lambda_{em}) I_{An}(\lambda_{ex}) e^{-\alpha(\lambda_{em})y} \quad (8)$$

$$I_{3D}(\lambda_{ex}, \lambda_{em}) = \frac{4I_0 e^{-\frac{\alpha(\lambda_{ex})l}{2}} \sinh\left(\frac{\alpha(\lambda_{ex})\Delta_x}{2}\right) e^{-\frac{\alpha(\lambda_{em})l}{2}} \sinh\left(\frac{\alpha(\lambda_{em})\Delta_y}{2}\right)}{\Delta_y \alpha(\lambda_{ex}) \alpha(\lambda_{em})} \sum_n \alpha_n(\lambda_{ex}) \Phi_n \gamma_n(\lambda_{em}) \quad (9)$$

153  $\frac{\alpha(\lambda_{ex})\Delta_x}{2}$  and  $\frac{\alpha(\lambda_{em})\Delta_y}{2}$  are supposed to be small enough to make the following  
 154 approximation:

$$I_{3D}(\lambda_{ex}, \lambda_{em}) = I_0 \Delta_x \left( \sum_n \alpha_n(\lambda_{ex}) \Phi_n \gamma_n(\lambda_{em}) \right) e^{-\frac{\alpha(\lambda_{ex})l}{2}} e^{-\frac{\alpha(\lambda_{em})l}{2}} \quad (10)$$

155 Then we can define  $g = \frac{l}{2}$  and  $G_n = I_0 \Delta_x \Phi_n$ . This leads to the final expression  
 156 of the model:

$$I_{3D}(\lambda_{ex}, \lambda_{em}) = \left( \sum_{n=1}^N G_n c_n \varepsilon_n(\lambda_{ex}) \gamma_n(\lambda_{em}) \right) \prod_{n=1}^N e^{-g(c_n \varepsilon_n(\lambda_{ex}) + c_n \varepsilon_n(\lambda_{em}))} \quad (11)$$

157 In the following we define

$$L(\lambda_{ex}, \lambda_{em}) = \sum_{n=1}^N G_n c_n \varepsilon_n(\lambda_{ex}) \gamma_n(\lambda_{em}) \quad (12)$$

158 then we have

$$I_{3D}(\lambda_{ex}, \lambda_{em}) = L e^{-g(c_n \varepsilon_n(\lambda_{ex}) + c_n \varepsilon_n(\lambda_{em}))} \quad (13)$$

159 This equation is clearly equivalent to equation 5 with  $I_c = L$ . Its first order  
 160 approximation is justified for small enough concentrations. In this case, since  
 161 the exponential term tends to 1 one obtains the linear model of fluorescence.  
 162 Correction of inner filter effects simplifies spectral analysis. It is interesting  
 163 to note that their modelization is also used in another context. Actually, a  
 164 recent article [28] highlighted the major contribution of inner filter effects in  
 165 the phenomenon of concentration-dependent red-shift [29,30]. In this work, a  
 166 similar model has been successfully used to optimize synchronous fluorescence  
 167 spectroscopy of concentrated mixtures of fluorophores.

169 The previous model describes non linear effects but the related equation can  
 170 still be considered as a bilinear decomposition involving some modified indi-  
 171 vidual spectra  $\varepsilon'_n(\lambda_{ex})$  and  $\gamma'_n(\lambda_{em})$  :

$$I_{3D}(\lambda_{ex}, \lambda_{em}) = \sum_{n=1}^N \varepsilon'_n(\lambda_{ex}) \gamma'_n(\lambda_{em}) \quad (14)$$

172 with,

$$\varepsilon'_n(\lambda_{ex}) = G_n c_n \varepsilon_n(\lambda_{ex}) e^{-\sum_{p=1}^N g c_p \varepsilon_p(\lambda_{ex})} \quad (15)$$

$$\gamma'_n(\lambda_{em}) = \gamma_n(\lambda_{em}) e^{-\sum_{p=1}^N g c_p \varepsilon_p(\lambda_{em})} \quad (16)$$

173 It is well known that bilinear decompositions have an infinite number of equiv-  
 174 alent solutions in the least square sense. Therefore, without additional infor-  
 175 mation, no mathematical tool can diagnose whether a FEEM is affected by  
 176 inner filter effects or not. *A fortiori* additional informations are also needed to  
 177 correct inner filter effects. In ACA this information is the solution absorbance  
 178 spectrum. This section shows how the correction can be made with fluorescent  
 179 spectra only.

180 According to equation 11 the FEEM  $I_{3D}$  of a  $N$  fluorophores mixture is the  
 181 product of a linear ( $L$ ) term in respect of concentrations and spectra by a non  
 182 linear one, denoted  $H$ :

$$H(\lambda_{ex}, \lambda_{em}) = \prod_{n=1}^N \exp(-g(c_n \varepsilon_n(\lambda_{ex}) + c_n \varepsilon_n(\lambda_{em}))) \quad (17)$$

183 So we can write:

$$I_{3D}(\lambda_{ex}, \lambda_{em}) = L(\lambda_{ex}, \lambda_{em}) H(\lambda_{ex}, \lambda_{em}) \quad (18)$$

184 Now, let's  $I_{3D_p}$  be the FEEM of the same mixture, diluted by a factor  $p$ , then  
 185 we have:

$$I_{3D_p}(\lambda_{ex}, \lambda_{em}) = \left( \sum_{n=1}^N G_n \frac{c_n}{p} \varepsilon_n(\lambda_{ex}) \gamma_n(\lambda_{em}) \right) \prod_{n=1}^N e^{-g(\frac{c_n}{p} \varepsilon_n(\lambda_{ex}) + \frac{c_n}{p} \varepsilon_n(\lambda_{em}))} \quad (19)$$

$$I_{3D_p}(\lambda_{ex}, \lambda_{em}) = \frac{1}{p} L(\lambda_{ex}, \lambda_{em}) H^{\frac{1}{p}}(\lambda_{ex}, \lambda_{em}) \quad (20)$$

186 The analytical resolution of 18 and 20 gives

$$L(\lambda_{ex}, \lambda_{em}) = \left( \frac{(p I_{3D_p}(\lambda_{ex}, \lambda_{em}))^p}{I_{3D}(\lambda_{ex}, \lambda_{em})} \right)^{\frac{1}{p-1}} \quad (21)$$

$$H(\lambda_{ex}, \lambda_{em}) = \left( \frac{I_{3D}(\lambda_{ex}, \lambda_{em})}{p I_{3D_p}(\lambda_{ex}, \lambda_{em})} \right)^{\frac{p}{p-1}} \quad (22)$$

The  $L$  term is the corrected FEEM estimated by CDA, corresponding to the linear model of fluorescence. As previously mentioned, the correction only requires the original FEEM and the diluted FEEM and the value of the dilution factor  $p$ .

The sensitivity of the estimator of  $L$  to  $p$  is difficult to quantify. A first order approximation of the variability of  $L$  ( $\Delta_L$ ) leads to:

$$\frac{\Delta_L}{L} = \frac{(p - 1 - \log(p) - \log(I_{3Dp}) + \log(I_{3D}))\Delta_p}{(p - 1)^2}$$

187 According to this equation, a high factor should be preferred. However it would  
 188 involve the drawbacks of a strong dilution (see section 1.2). Finally, we advo-  
 189 cate for a dilution factor corresponding to the simplest dilution process, thus  
 190 the experimental uncertainty  $\Delta_p$  is minimized. This was the case for all the  
 191 experiments presented in this study.

192 Owing to the term by term division in equation 21, noisy values in the mea-  
 193 sured FEEM could affect the estimation of  $L$ . Actually if the division involves  
 194 two small values relatively to the noise level, some very narrow and localized  
 195 peaks can appear. Fluorescent spectroscopy is a very sensitive technique there-  
 196 fore this kind of deviation are rarely observed in practical situations. Otherwise  
 197 those peaks appear outside the main fluorescing areas. In consequence, they  
 198 can be easily detected and filtered without damaging the fluorescent peaks.

199

### 200 2.3 CDA and multilinear analysis of concentrated fluorescing mixtures

201 We consider now a set of  $I$  mixtures and  $c_n(i)$  denotes the concentration of  
 202 fluorophore  $n$  in mixture  $i$ . CDA methodology is simple, the correction is done  
 203 sample by sample. The first step consists in choosing the dilution factor  $p$  for  
 204 each sample (see the end of section 2.2). Obviously the same value can be used  
 205 for every sample. Then, the corresponding controlled dilution is performed  
 206 and both FEEM  $I_{3D}$  and  $I_{3Dp}$  are measured. Before correction, Rayleigh and  
 207 Raman scatters must be corrected carefully on each FEEM. This is the end  
 208 of the experimental and pre-processing steps.

209 Finally, for each sample  $i$ , the estimation  $L(i, \lambda_{ex}, \lambda_{em})$  of the linearized FEEM  
 210 is obtained directly from equation 21. At this stage, the correction of the inner  
 211 filter effects is completed.

212 Actually we have to take into account measurement and modelization errors.  
 213 In addition we can define  $\tilde{c}_n(i) = G_n c_n(i)$ . Therefore in practice, definition 12  
 214 is rewritten:

$$L(i, \lambda_{ex}, \lambda_{em}) = \sum_{n=1}^N \tilde{c}_n(i) \varepsilon_n(\lambda_{ex}) \gamma_n(\lambda_{em}) + E(i, \lambda_{ex}, \lambda_{em}) \quad (23)$$

215 where  $E$  is the error term. Equation 23 is a rank  $N$  decomposition of the 3  
 216 way tensor  $L$  or in other words a 3-way PARAFAC model of rank  $N$ . For  
 217 each fluorophore  $n$ , the loading vectors of the decomposition  $\tilde{c}_n$ ,  $\varepsilon_n$ , and  $\gamma_n$   
 218 are linearly linked to its concentration profile, its excitation spectrum and its  
 219 emission spectrum respectively. Moreover, the solution of this decomposition  
 220 is unique up to trivial scaling and position indeterminacy [31,32]. Finally, sev-  
 221 eral efficient algorithms were proposed and compared for the estimation of the  
 222 loading vectors. These are largely described in the literature [33–35]. Those  
 223 three physical, mathematical and practical reasons made the PARAFAC de-  
 224 composition the most suitable tool for analysing linear(ized) FEEM. Tutorials  
 225 and examples of PARAFAC application to FEEM analysis can be found else-  
 226 where [36,22,37].

227 Eventually, the PARAFAC decomposition can be run normally on the cor-  
 228 rected FEEM set in order to find out real individual spectra and concentration  
 229 profiles of each fluorophore.

230

### 231 3 Experimental

#### 232 3.1 Data set 1, standard mixtures

233 Seven solutions ( $S_1^i, i = 1 \dots 7$ ) with different concentrations of fluorescein  
 234 (Aldrich) and quinine sulphate (Merck) were prepared in 0.1M  $H_2SO_4$  (Aldrich)  
 235 in order to validate the correction method. All chemicals are analytical grade.  
 236 Concentrations in fluorescein and quinine sulphate are given in table 1 along  
 237 with solution absorbances. These two fluorophores and their concentrations in  
 238 the solutions were chosen because of their good fluorescing ability and their  
 239 overlapped spectra in order to emphasize inner filter effects.

Table 1

Concentrations, maximal absorbances and mean absorbances of the original solu-  
 tions of quinine sulphate and fluorescein.

Solution	$S_1^1$	$S_1^2$	$S_1^3$	$S_1^4$	$S_1^5$	$S_1^6$	$S_1^7$
$c_{SQ}$ (ppm)	0	11.02	32.6	54.38	76.15	97.73	108.75
$c_F$ (ppm)	83.15	74.72	58.23	41.58	24.92	8.43	0
Absorbance max.	2.30	2.18	1.83	1.32	1.17	1.38	1.47
Mean Absorbance.	0.42	0.40	0.37	0.33	0.29	0.27	0.24

Concentrations in quinine sulphate ( $c_{SQ}$ ) and fluorescein ( $c_F$ ) are given in parts per  
 million (ppm). Maximum and mean value of absorbance are relative to the 275 to  
 500 nm excitation range.

240 Seven twice diluted solutions ( $S_{1D}^i, i = 1 \dots 7$ ) were obtained by mixing equal  
 241 volumes of initial solutions  $S_1^i$  and 0.1M of  $H_2SO_4$ . Table 2 gives the actual  
 242 value of the dilution factor for the seven solutions and the standard deviation  
 243 due to the pipet precision.

Table 2

Dilution factors used for the seven solutions of fluorescein and quinine sulphate.

Solution	$S_{1D}^1$	$S_{1D}^2$	$S_{1D}^3$	$S_{1D}^4$	$S_{1D}^5$	$S_{1D}^6$	$S_{1D}^7$
$p$	2.11	2.08	2.07	2.07	2.07	2.08	2.11
$\sigma_p$	0.034	0.032	0.030	0.029	0.030	0.032	0.034

$\sigma_p$  is the estimated standard deviation of the dilution factor due to the experimental dilution

244 Reference solutions ( $S_{1R}^i, i = 1 \dots 7$ ) were obtained by diluting 100  $\mu$ L of  $S_1^i$   
 245 in 3000  $\mu$ L of 0.1M of  $H_2SO_4$ . In this case of simple mixtures, this dilution  
 246 prevents inner filter effects without physico-chemical changes.

247 All measured spectra were obtained with a fluorometer Hitachi F4500. FEEM  
 248 of the three solutions sets  $S_1, S_{1D}$  and  $S_{1R}$ , were recorded at 30000 nm/min  
 249 scan speed from 350 to 700 nm in emission by step of 5 nm and for excitation  
 250 wavelength from 275 to 500 nm by step of 5 nm. Excitation and emission band-  
 251 width were 5 nm. Fluorescence intensity was corrected from PM response using  
 252 manufacturer setting. Data for FEEM treatment were extracted by FLWinLab  
 253 software for emission and excitation range stepped every 5 nm. Rayleigh and  
 254 Raman scatters were removed numerically by the method proposed by Zepp  
 255 in [38]. In the following, measured FEEM from original, diluted and reference  
 256 solution  $i$  will be referred as  $I_{1U}^i, I_{1D}^i$  and  $I_{1R}^i$  respectively. Absorption spec-  
 257 tra of solutions  $S_1$  were obtained from transmittance spectra recorded with  
 258 absorbance mode of the F4500 (speedscan 240 nm/min) from 200 to 800 nm  
 259 with 5 nm bandwidths in excitation and emission. 2D reference spectra of flu-  
 260 orescein and quinine sulphate were recorded from  $S_{1R}^1$  and  $S_{1R}^7$  respectively, at  
 261 240 nm/min scan speed by step of 1 nm with 5 nm bandwidth in excitation  
 262 and 2.5 nm bandwidth in emission. Quinine sulphate ( $I_{SQ-ex}$ ) and fluorescein  
 263 ( $I_{F-ex}$ ) excitation spectra were recorded from 275 to 500 nm at 450 nm and  
 264 510 nm emission wavelength respectively. Their emission spectra ( $I_{SQ-em}$  and  
 265  $I_{F-em}$ ) were recorded from 350 to 700 nm at 340 nm and 440 nm excitation  
 266 wavelength respectively.

267  $I_{1R}$  and  $I_{1U}$  compose the groups of reference and uncorrected FEEM respec-  
 268 tively. Using absorption spectra in equation 5, we could compute the ACA  
 269 corrected FEEM ( $I_{1ACA}^i$ ) from  $I_{1U}^i, i = 1 \dots 7$ . In the same way, using  $I_{1D}^i$  in  
 270 equation 21 with dilution factor values of table 2, we could apply CDA to  $I_{1U}^i$   
 271 and compute the CDA corrected FEEM ( $I_{1CDA}^i$ ),  $i = 1 \dots 7$ .  $I_{1ACA}$  and  $I_{1CDA}$   
 272 compose the groups of ACA and CDA corrected FEEM respectively.

273 These four groups of FEEM are considered as four 3-way tensors of dimensions  
 274  $7 \times 46 \times 51$ . Trilinear decompositions of these tensors were performed with the  
 275 PARAFAC-ALS algorithm of the nway toolbox for Matlab [39].

277 The second data set is composed of FEEM obtained from eleven samples  
 278 of concentrated humic acid solutions which were extracted from catchments  
 279 of Cameron soils. For each sample, 1 g of soil was extracted by 30 mL of  
 280  $HCl$  (1M) solution. After separation, the supernatant solution was cleaned on  
 281 XAD-8 resin and stored at  $4^\circ C$  in dark. The resting soil was then extracted  
 282 with 30 mL of 1M  $NaOH$  solution. After separation, this second supernatant  
 283 contains humic acid substance. Purification was done by acidic precipitation  
 284 and sodic redissolution. Humic acid gave dark brown solution and fulvic acid  
 285 yellow solution. Original solutions ( $S_2^i, i = 1 \dots 11$ ), were obtained by dilut-  
 286 ing 100  $\mu L$  of the extracted solutions in 3000  $\mu L$  of 0.1M  $NaOH$  buffer. All  
 287 chemicals are analytical grade. Eleven twice diluted solutions in 0.1M  $NaOH$   
 288 buffer ( $S_{2D}^i, i = 1 \dots 11$ ) were also prepared. Finally, reference solutions with-  
 289 out inner filter effect ( $S_{2R}^i, i = 1 \dots 11$ ) were obtained by 15 times dilutions of  
 290  $S_{2D}^i$ .

291 FEEM of  $S_2^i$  ( $I_{2U}^i$ ),  $S_{2D}^i$  ( $I_{2D}^i$ ) and  $S_{2R}^i$  ( $I_{2R}^i$ ) were recorded for  $i = 1 \dots 11$ , at  
 292 30000 nm/min scan speed from 340 to 650 nm in emission for excitation wave-  
 293 length from 240 to 600 nm, by emission and excitation step of 5 nm and with  
 294 10 nm bandwidth in excitation and 5 nm bandwidth in emission. Rayleigh  
 295 and Raman scatters were consistently corrected [38].

296  $I_{2R}$  and  $I_{2U}$  compose the groups of reference and uncorrected FEEM respec-  
 297 tively. Using  $I_{2D}^i$  in equation 21 with  $p = 2$ , we could apply CDA to  $I_{2U}^i$  and  
 298 compute the CDA corrected FEEM ( $I_{2CDA}^i$ ),  $i = 1 \dots 11$ .  $I_{2CDA}$  compose the  
 299 group of CDA corrected FEEM. Comparison with ACA was not made on this  
 300 data set.

301 Hence, three tensors of dimensions  $11 \times 73 \times 63$  were obtained and decomposed  
 302 by PARAFAC-ALS [39].

## 303 4 Results and discussion

304 In the following the PARAFAC applications to uncorrected, reference, ACA  
 305 corrected and CDA corrected groups of FEEM will be referred as U-PARAFAC,  
 306 R-PARAFAC, ACA-PARAFAC and CDA-PARAFAC respectively.

### 307 4.1 Data set 1

308 Each CDA corrected FEEM ( $I_{1CDA}^i, i = 1 \dots 7$ ) are firstly compared to  $I_{1R}^i$ ,  
 309  $I_{1U}^i$  and  $I_{1ACA}^i$ . Representative examples of these FEEM are presented on figure  
 310 6 for  $i = \{1, 3, 4, 6\}$ . In these examples, CDA provides satisfying estimations

311 of the reference FEEM in spite of a small distortion in the fluorescein peak  
 312 (440/510nm), particularly on  $I_{1CDA}^1$ . In order to quantify these comparisons  
 313 three relative squared residual error terms ( $r_{1CDA}^i$ ,  $r_{1U}^i$  and  $r_{1ACA}^i$ ) were com-  
 314 puted as follow for each solution  $i$  and stored in table 3.

$$r_{1CDA}^i = \frac{\sum_{j,k}(I_{1R}^i(j,k) - I_{1CDA}^i(j,k))^2}{\sum_{i,j} I_{1R}^i(j,k)^2} \quad (24)$$

$$r_{1U}^i = \frac{\sum_{j,k}(I_{1R}^i(j,k) - I_{1U}^i(j,k))^2}{\sum_{j,k} I_{1R}^i(j,k)^2} \quad (25)$$

$$r_{1ACA}^i = \frac{\sum_{j,k}(I_{1R}^i(j,k) - I_{1ACA}^i(j,k))^2}{\sum_{j,k} I_{1R}^i(j,k)^2} \quad (26)$$

The significance of the non linear term in equation (11) increases with the

Table 3

Comparison of the relative squared residual error terms(in % ) for the seven solution of data set 1.

Solution	1	2	3	4	5	6	7
$r_{1U}$	117	109	81	46	29	20	13
$r_{1ACA}$	31	38	59	33	14	7	8
$r_{1CDA}$	11	5	2	1.8	0.3	0.06	0.03

315 solution absorbance. Therefore comparing tables 1 and 3, there is an obvious  
 316 correlation between  $r_{1U}$  and the mean absorbances as both values decrease  
 317 regularly from solutions 1 to 7. The same observation holds true regarding  
 318  $r_{1CDA}$ . The correlation is less apparent for  $r_{1ACA}$  but globally the error is  
 319 greater for solutions 1 to 4 than for the least absorbing solutions.  $r_{1U}$  values  
 320 are comprised between 13% and 117%. After CDA correction, these boundary  
 321 values decreased to 0.03% and 11% respectively . These results are very sat-  
 322 isfying for solutions 2 to 7. The first solution shows a stronger error but there  
 323 is still a clear improvement in comparison of the original FEEM. CDA results  
 324 are always clearly better than ACA ones. Actually, in the more favourable  
 325 case in respect to ACA results (solution 1),  $r_{1ACA}$  is almost 3 times greater  
 326 than  $r_{1CDA}$ .

327 These results showed that CDA provided a better estimation of  $I_{1R}$  than  
 328 ACA. In order to verify if CDA correction is satisfying for further analysis,  
 329 the PARAFAC-ALS algorithm was applied to the four groups of FEEM. For  
 330 each group, the core consistency diagnostic (CORCONDIA) [40] suggested  
 331 two as the right number of components. Consequently, each PARAFAC de-  
 332 composition provides an estimation of the quinine sulphate and fluorescein  
 333 excitation and emission spectra and an estimation of their relative concen-  
 334 trations trough the data set. After normalisation, the relative squared error  
 335 ( $r^c$ ,  $r^{ex}$  and  $r^{em}$ ) between the PARAFAC loadings and the real variables are  
 336 compared on tables 4 to 6. In the case of the spectral loadings, two other spec-  
 337 troscopic criteria are also used: the relative error to the maximum value of the  
 338

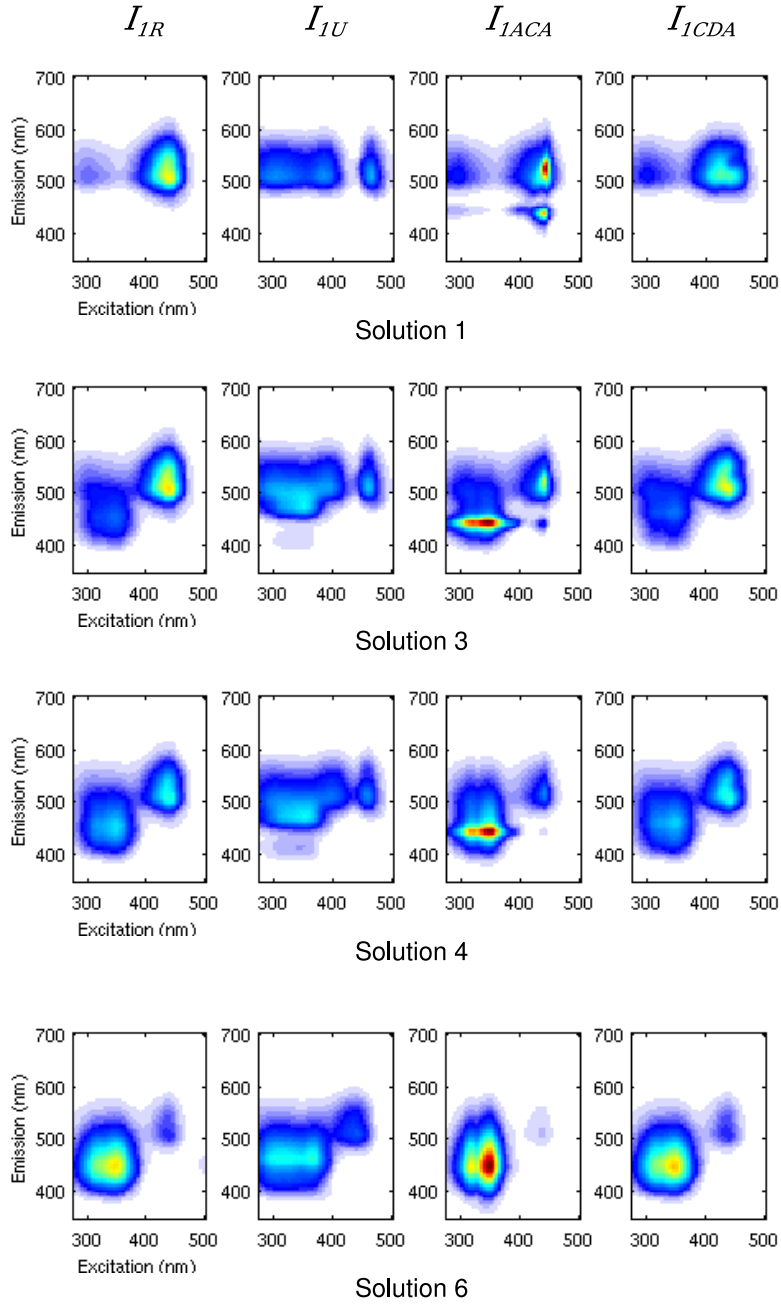


Fig. 6. Illustrations of  $I_{IR}$ ,  $I_{IU}$ ,  $I_{IACA}$  and  $I_{ICDA}$  for solutions 1, 3, 4 and 6 of data set 1.

339 spectrum and the shift on the position of the maximum. In addition, load-  
 340 ings obtained from U-PARAFAC and CDA-PARAFAC are shown on figures  
 341 7 to 9, along with real spectra and profiles. A first global remark should be  
 342 made: the perfect agreement on the three modes between the real variables and  
 343 R-PARAFAC loadings demonstrates that estimation errors of U-PARAFAC,



344 ACA-PARAFAC and CDA-PARAFAC are mainly due to inner filter effects  
 345 and not to the PARAFAC decomposition.

346

Table 4

Data set 1, concentration mode results. Relative squared residual error in percent between the real profiles and their estimation from  $I_{1R}$  ( $r_{1R}^c$ ),  $I_{1U}$  ( $r_{1U}^c$ ),  $I_{1ACA}$  ( $r_{1ACA}^c$ ) and  $I_{1CDA}$  ( $r_{1CDA}^c$ ).

Fluorophore	$r_{1R}^c$	$r_{1U}^c$	$r_{1ACA}^c$	$r_{1CDA}^c$
Quinine sulphate	0.1	8.3	0.7	0.2
Fluorescein	0.08	23	0.4	4.2

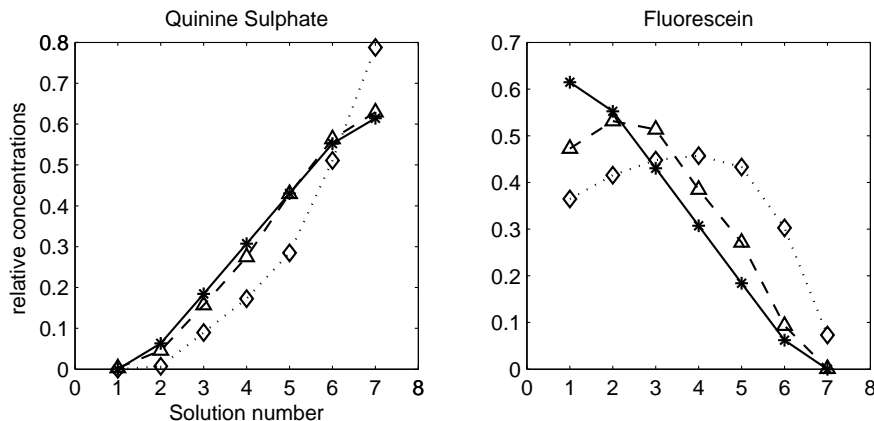


Fig. 7. Data set 1, PARAFAC loadings of the concentration mode: Real spectra (solid \* line), U-PARAFAC loadings (dot ◇ line) and CDA-PARAFAC loadings (dash △ line)

347 Results for the concentration mode are presented in table 4 and figure 7.  
 348 The shape of the concentration profile of the quinine sulphate is slightly af-  
 349 fected by filter effects ( $r_{1U}^c = 8.3\%$ ). However CDA-PARAFAC gives much  
 350 more accurate results for each solution ( $r_{1CDA}^c = 0.2\%$ ). ACA-PARAFAC is  
 351 also satisfying but it is not as efficient as CDA. The concentration profile of  
 352 the fluorescein is more distorted ( $r_{1U}^c = 23\%$ ). CDA-PARAFAC performs well  
 353 ( $r_{1CDA}^c = 4.2\%$ ) but the relative error is still high for the first solution. It  
 354 should be noted that ACA-PARAFAC ( $r_{1ACA}^c = 0.4\%$ ) do better than CDA-  
 355 PARAFAC. This last result is surprising because it is in contradiction with  
 356 the five other loadings.

357

358 Results for the excitation mode are presented in table 5 and figure 8. The  
 359 excitation mode is the most affected by inner filter effects, with  $r_{1U}^{ex}$  values of  
 360 12.9% and 95.15% for quinine sulphate and fluorescein respectively. Consid-  
 361 ering figure 8, quinine sulphate excitation spectrum is widened and flattened  
 362 by inner filter effects. These distortions are completely eliminated by CDA-

Table 5

Data set 1, excitation mode results. Comparison with the real spectra among three criteria: relative error on the maximum value (%), shift (nm) and relative squared error (%)

Fluorophore	Criterion	Ref.	Unc.	ACA cor.	CDA cor.
Quinine sulphate	Rel. err. max. val.	0.4	24	23	0.5
	Shift.	0	20	0	0
	$r^{ex}$	0.06	12.9	4.1	0.08
Fluorescein	Rel. err. max. val.	2.5	50.5	15	17
	Shift	0	20	5	5
	$r^{ex}$	0.07	95.15	9.3	3.4

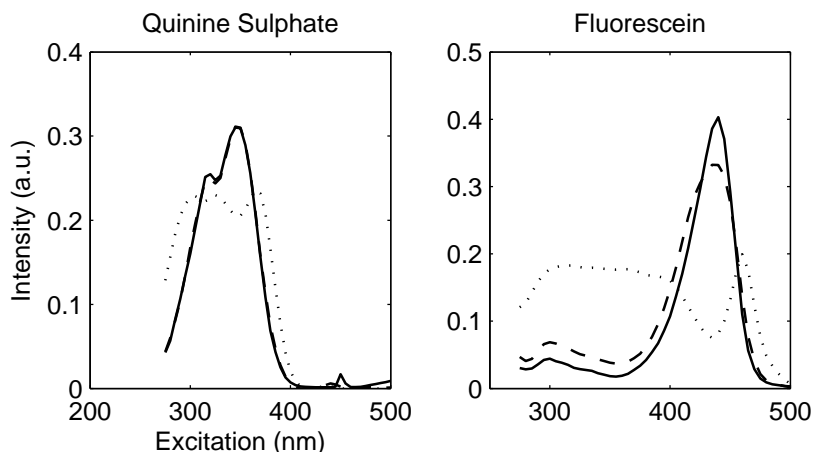


Fig. 8. Data set 1, PARAFAC loadings of the excitation mode: Real spectra (solid line), U-PARAFAC loadings (dot line) and CDA-PARAFAC loadings (dash line)

363 PARAFAC. Indeed, the CDA-PARAFAC estimated spectrum is really close  
 364 to the real spectrum ( $r_{ICDA}^{ex} = 0.08\%$ ). ACA-PARAFAC ( $r_{IACA}^{ex} = 4.1\%$ ) glob-  
 365 ally improves U-PARAFAC result ( $r_{IU}^{ex} = 12.9\%$ ). The relative error on the  
 366 maximum value is negligible with CDA-PARAFAC (0.5%) while it is impor-  
 367 tant with ACA-PARAFAC (23%) as without correction (24%). The 20 nm  
 368 shift on the position of the maximum is perfectly corrected by both CDA-  
 369 PARAFAC and ACA-PARAFAC. These observations hold true for the fluo-  
 370 rescein spectrum but the gaps between the different estimations are wider.  
 371 The spectrum estimated by U-PARAFAC is totally distorted in respect to the  
 372 real spectrum,  $r_{IU}^{ex} = 95.15\%$ . In spite of some residual distortions (flattening  
 373 and widening), the spectrum shape is almost fully recovered with CDA  
 374 correction ( $r_{ICDA}^{ex} = 3.4\%$ ). ACA-PARAFAC global estimation is not as good  
 375 ( $r_{IACA}^{ex} = 9.3\%$ ). In absence of correction, the relative error on the maximum  
 376 value is very high (50.5%). ACA-PARAFAC and CDA-PARAFAC relative  
 377 error are equivalent with 15% and 17% respectively. In the same way the

378 maximum shift is limited by ACA-PARAFAC and CDA-PARAFAC to the  
 379 fluorometer excitation step (5 nm) against 20 nm with U-PARAFAC.

Table 6

Data set 1, emission mode results. Comparison with the real spectra among three criteria: relative error on the maximum value (%), shift (nm) and relative squared error (%)

Fluorophore	Criterion	Ref.	Unc.	ACA cor.	CDA cor.
Quinine sulphate	Rel. err. max. val.	1.4	4.1	4.1	0.7
	Shift	0	5	5	0
	$r^{em}$	0.06	1.8	0.17	0.1
Fluorescein	Rel. err. max. val.	3.9	4	4.7	0.6
	Shift	0	0	15	0
	$r^{em}$	0.4	7.6	15	0.5

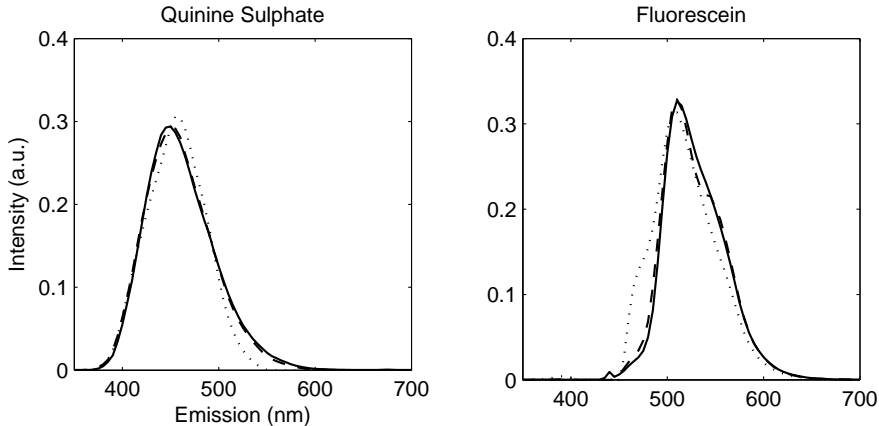


Fig. 9. Data set 1, PARAFAC loadings of the emission mode: Real spectra (solid line), U-PARAFAC loadings (dot line) and CDA-PARAFAC loadings (dash line)

380 Results for emission mode are presented in table 6 and figure 9. The emis-  
 381 sion spectrum of quinine sulphate is well estimated by U-PARAFAC ( $r_{1U}^{em} =$   
 382 1.8%). However ACA-PARAFAC and CDA-PARAFAC still improve this re-  
 383 sult in different proportions ( $r_{1ACA}^{em} = 0.17\%$  and  $r_{1CDA}^{em} = 0.1\%$ ). One should  
 384 note that the relative error on the maximum value is the same with ACA-  
 385 PARAFAC than with U-PARAFAC (4.1%). This small error is corrected by  
 386 CDA-PARAFAC (0.7%). In the same way, The 5nm shift is only corrected  
 387 by CDA-PARAFAC. A closer look should be given to fluorescein spectrum.  
 388 This for two reasons: Firstly, CDA-PARAFAC removes the shift of the right  
 389 slope and corrects the large distortion on the left slope of the peak (figure 9).  
 390 It also creates a weaker distortion on the right slope. This distortion is neg-  
 391 ligible but it is very similar to the spectral distortion observed on the CDA  
 392 corrected FEEM of solution 1 (figure 6). Secondly, ACA-PARAFAC shows its

393 real limitation on this loading as the estimated spectrum is worse than the one  
394 estimated by U-PARAFAC regarding the three criteria (table 6). The relative  
395 squared error is twice higher (15% against 7.6%) with ACA-PARAFAC than  
396 with U-PARAFAC while it is negligible with CDA-PARAFAC (0.5%). The  
397 relative error on the maximum value is quite small with U-PARAFAC (4%)  
398 and ACA-PARAFAC (4.7%) but it is totally corrected by CDA-PARAFAC  
399 (0.7%). Eventually, ACA-PARAFAC introduces a 15 nm shift on the position  
400 of the maximum which do not exist with U-PARAFAC and CDA-PARAFAC.  
401 In conclusion to this test, as expected, U-PARAFAC provides the worst re-  
402 sults. These are very bad, specially for the fluorescein excitation spectrum and  
403 quinine sulphate emission spectrum whose the wavelength domains strongly  
404 overlap. ACA-PARAFAC improves these results. Regarding ACA results, load-  
405 ings estimation is better than expected. However, it is outperformed by CDA-  
406 PARAFAC at the exception of the fluorescein concentration profile. CDA-  
407 PARAFAC results are indeed closer to those obtained with the reference  
408 FEEM although a larger error is observed on fluorescein excitation (table  
409 5) and concentration loadings (table 6). This must be seen as the PARAFAC  
410 manifestation of the small distortion observed on the CDA correction of so-  
411 lution 1 (figure 6) and more generally this is an indication of the CDA lim-  
412 itations. Actually, CDA is limited to a certain domain of validation because  
413 equation 11 is obtained after several approximations. This holds true for any  
414 correction method relied on equation 11. Regarding the results of table 3 and  
415 figure 7, this limitation has probably been reached with solution 1. In these  
416 cases of very high absorbance (equal or above 2), more sophisticated models  
417 should be used. Nevertheless, we have demonstrated here that the validation  
418 domain of CDA is much larger than the linear one. Then the FEEM provided  
419 by CDA are close enough to the ideal linear FEEM to allow advanced spectral  
420 data analysis such as the PARAFAC decomposition while this is not the case  
421 with uncorrected FEEM or with ACA to a lesser extent. This example also  
422 show that CDA-PARAFAC improves both kinds of PARAFAC results: On the  
423 one hand it allows to recover the overall profile of strongly distorted loadings,  
424 on the other hand it provides some very accurate estimations of less affected  
425 loadings.

#### 426 4.2 *Data set 2, application to field*

427 In this section, performances of CDA and CDA-PARAFAC are shown for the  
428 correction and the decomposition of mixtures of model molecules. The first  
429 stage of the test is a comparison between the reference FEEM ( $I_{2R}$ ), the un-  
430 corrected FEEM ( $I_{2U}$ ) and the CDA corrected FEEM ( $I_{2CDA}$ ) of data set 2.  
431 Four representative examples of these different FEEM are presented on figure  
432 10. Regarding these examples a main distortion appears on the fluorescence  
433 pattern if no correction is applied. It turns out that filter effects increase the

434 peak around (480 nm, 520 nm) and decrease the intensity of the fluorescence  
 435 signal located under 400 nm in excitation. This area is composed by several  
 436 peaks which are partially recovered by CDA. The PARAFAC decompositions  
 437 should show whether this finer aspect of the correction is satisfying or not.  
 438 The relative squared residual error terms ( $r_{2U}$  and  $r_{2CDA}$ ) of the whole data  
 439 set are given in table 7. FEEM obtained with CDA are not as closed to the  
 440 reference FEEM as for data set 1. Actually,  $r_{2CDA}$  average value is about 4%  
 441 while it is about 55% for  $r_{2U}$ . Hence, it proves that CDA correction is still  
 442 very beneficial relatively to the uncorrected FEEM.

443

Table 7

Comparison of the relative squared residual error terms (in %) for the eleven solution of data set 2

Sample	1	2	3	4	5	6	7	8	9	10	11
$r_{2U}$	51	66	35	59	35	44	83	76	83	10	60
$r_{2CDA}$	3	3	5	1.4	5	3	4	5	6	5	2

444 The PARAFAC-ALS algorithm was applied to  $I_{2R}$ ,  $I_{2U}$  and  $I_{2CDA}$ . In oppo-  
 445 sition to set 1, the real number of fluorophore is unknown. This is actually  
 446 the main problem with PARAFAC analysis of unknown FEEM. We com-  
 447 pared the results provided by three classical tests: residual variance analysis,  
 448 split half analysis and CORCONDIA. Finally, three components were used for  
 449 the decompositions. These will be labelled fluorophore 1, 2 and 3 in the fol-  
 450 lowing. The real fluorophores are also unknown, consequently, R-PARAFAC  
 451 loadings are taken as references for the evaluation of U-PARAFAC and CDA-  
 452 PARAFAC results. This comparison is made on tables 8 to 10 and figures 11  
 453 to 13.

454

Table 8

Data set 2, concentration mode results. Relative squared residual error in percent between the reference loadings and U-PARAFAC loadings ( $r_{2U}^c$ ) or CDA-PARAFAC loadings ( $r_{2CDA}^c$ ).

Fluorophore	$r_{2U}^c$	$r_{2CDA}^c$
1	6.6	5.6
2	26	4
3	7.3	8.2

455 Results for the concentration mode are presented in table 8 and figure 11. In-  
 456 ner filters and CDA correction have little effects on the profile of fluorophore  
 457 1 ( $r_{2U}^c = 6.6\%$  and  $r_{2CDA}^c = 5.6\%$ ) and 3 ( $r_{2U}^c = 7.3\%$  and  $r_{2CDA}^c = 8.2\%$ ). On  
 458 the opposite, fluorophore 2 is far more affected. The concentration profile esti-  
 459 mated by U-PARAFAC is clearly unsatisfying ( $r_{2U}^c = 26\%$ ). CDA-PARAFAC  
 460 provides an acceptable estimation as the error term decreases to 4 % . Finally

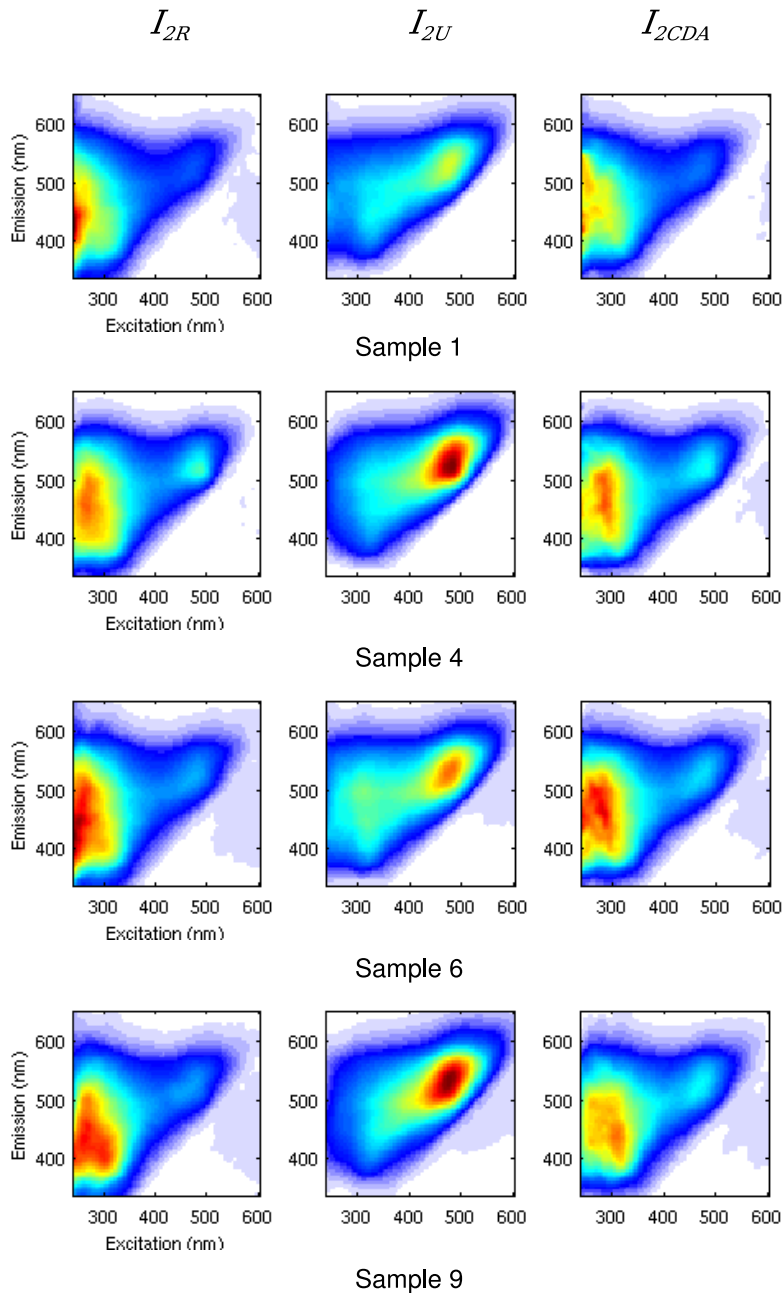


Fig. 10. Illustrations of  $I_{2R}$ ,  $I_{2U}$  and  $I_{2CDA}$  for samples 1, 4, 6 and 9 of data set 2.

461 the three estimated profiles by CDA-PARAFAC are satisfying at the exception  
 462 of sample 8 and 9. At the opposite of  $r_{2U}^c$ ,  $r_{2CDA}^c$  is greater for fluorophores 1  
 463 and 3. This is mainly due to the larger estimation errors on samples 8 and 9.  
 464

465 Results for excitation mode are presented in table 9 and figure 12. Excita-  
 466 tion mode is the most affected mode of the decomposition as for data set 1.  
 467 Estimation of fluorophore 1 excitation spectrum takes a clear advantage of

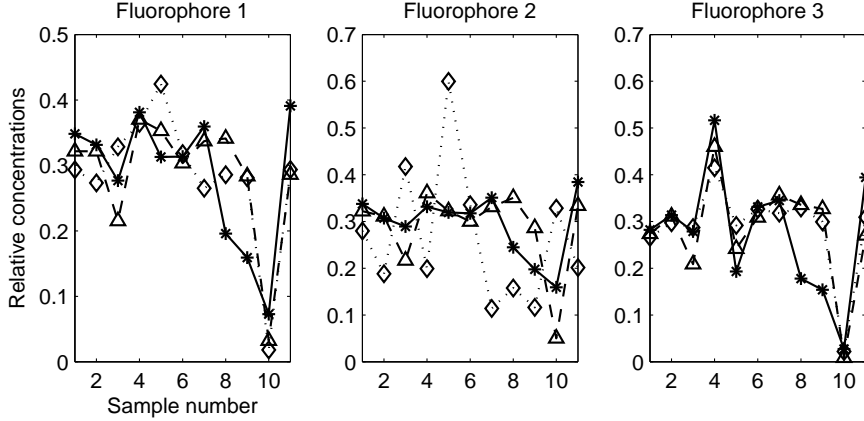


Fig. 11. Data set 2, PARAFAC loadings of the concentration mode: R-PARAFAC loadings (solid \* line), U-PARAFAC loadings (dot  $\diamond$  line) and CDA-PARAFAC loadings (dash  $\triangle$  line)

Table 9

Data set 2, excitation mode results. Comparison with the reference loadings among three criteria: relative error on the maximum value (%), shift (nm) and relative squared error (%)

Fluorophore	Test	Unc.	CDA cor.
1	Rel. err. max. val.	18	11
	Shift	185	15
	$r^{ex}$	110	1.5
2	Rel. err. max. val.	30	20
	Shift	20	0
	$r^{ex}$	9.6	2.1
3	Rel. err. max. val.	27.5	0.2
	Shift	10	5
	$r^{ex}$	35	0.6

468 CDA-PARAFAC. Regarding shape and position, the spectrum estimated by  
469 U-PARAFAC is a far cry from the reference spectrum ( $r_{2U}^{ex} = 110\%$ ). On the  
470 opposite, CDA-PARAFAC provides a very good estimation ( $r_{2CDA}^{ex} = 1.5\%$ ).  
471 The relative error on the maximum value is smaller (11 % against 18 %)  
472 but above all, the large shift (185 nm) is brought back to 15 nm. Despite U-  
473 PARAFAC estimation of fluorophore 2 spectrum is acceptable ( $r_{2U}^{ex} = 9.6\%$ ),  
474 CDA-PARAFAC improves this result ( $r_{2CDA}^{ex} = 2.1\%$ ). In absence of correc-  
475 tion, the double peak disappears. CDA-PARAFAC correctly restores this fea-  
476 ture but the relative error on the maximum value remains high (20%). On  
477 the other hand, the 20 nm shift is completely corrected. U-PARAFAC estima-

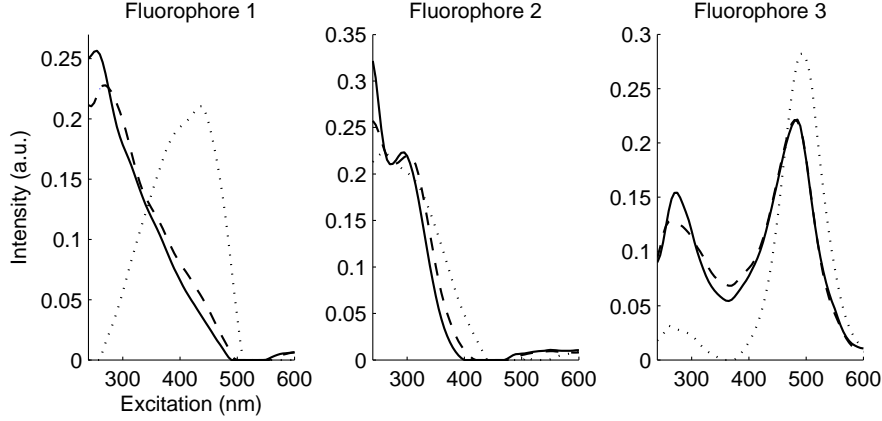


Fig. 12. Data set 2, PARAFAC loadings of the excitation mode: R-PARAFAC loadings (solid line), U-PARAFAC loadings (dot line) and CDA-PARAFAC loadings (dash line)

478 tion of fluorophore 3 spectrum is not satisfying ( $r_{2U}^{ex} = 35\%$ ). One of the two  
 479 peaks almost disappears while the second one is overestimated. Nevertheless,  
 480 its estimation by CDA-PARAFAC is almost identical to the reference spec-  
 481 trum ( $r_{2CDA}^{ex} = 0.6\%$ ). The relative error on the maximum value is significant  
 482 with U-PARAFAC (27.5%) but it becomes negligible with CDA-PARAFAC  
 483 (0.2%). The shift is also reduced from 10 nm to 5 nm.

Table 10

Data set 2, emission mode results. Comparison with the real spectra among Comparison with the reference loadings among three criteria: relative error on the maximum value (%), shift (nm) and relative squared error (%)

Fluorophore	Test	Orig.	Corr.Dil.
	Rel. err. max. val.	13	12
1	Shift	40	25
	$r^{em}$	21	4
	Rel. err. max. val.	21	1.8
2	Shift	60	10
	$r^{em}$	63	6
	Rel. err. max. val.	5.2	1.4
3	Shift	20	10
	$r^{em}$	8.8	0.9

484 Results for emission mode are presented in table 10 and figure 13. Estimation  
 485 of fluorophore 1 emission spectrum by U-PARAFAC is mitigated ( $r_{2U}^{em} = 21\%$ ).  
 486 CDA-PARAFAC overall result is very acceptable ( $r_{2CDA}^{em} = 4\%$ ). The spec-  
 487 trum is overestimated by both U-PARAFAC (relative error to the maximum



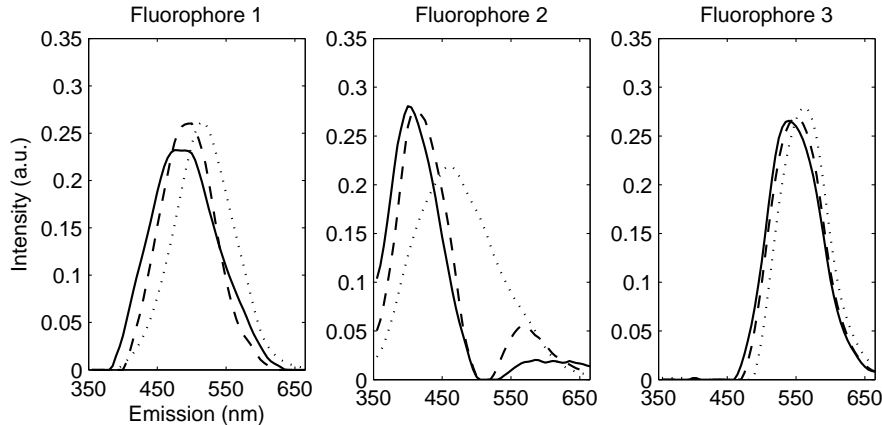


Fig. 13. Data set 2, PARAFAC loadings of the emission mode: R-PARAFAC loadings (solid line), U-PARAFAC loadings (dot line) and CDA-PARAFAC loadings (dash line)

488 value of 13%) and CDA-PARAFAC (12%). On the opposite, CDA-PARAFAC  
 489 limits to 25 nm the large shift (40 nm) observed when no correction is ap-  
 490 plied. Fluorophore 2 spectrum is more severely affected by inner filter ef-  
 491 fects. CDA-PARAFAC provides a very satisfying estimation of this spectrum  
 492 ( $r_{2CDA}^{em} = 6\%$ ) in respect to U-PARAFAC result ( $r_{2U}^{em} = 63\%$ ). The relative  
 493 error to the maximum value observed with U-PARAFAC is important (21%)  
 494 but it is well corrected by CDA-PARAFAC (1.8%). In the same way, the 60  
 495 nm shift is limited to 10 nm. Fluorophore 3 spectrum is correctly estimated  
 496 by U-PARAFAC ( $r_{2U}^{em} = 8.8\%$ ). CDA-PARAFAC still improves the estima-  
 497 tion ( $r_{2CDA}^{em} = 0.9\%$ ). The relative error to the maximum value is lower (1.4%  
 498 against 5.2%) and the shift is reduced from 20 nm to 10 nm.

499 CDA-PARAFAC results on data set 2 are also conclusive. All the loading are  
 500 indeed correctly estimated. Relative concentrations of only one fluorophore in  
 501 only two samples out of eleven are poorly estimated and all the estimated spec-  
 502 tra are close enough to the corresponding reference spectra. The crucial point  
 503 on this second example is that in a real case situation of DOM tracing, CDA-  
 504 PARAFAC would probably give interpretative results while the uncorrected  
 505 FEEM would provide misleading estimations. Moreover, excitation spectra of  
 506 fluorophores 1 and 3 illustrate the two main kinds of spectral deviations due  
 507 to inner filters. Fluorophore 1 shape is distorted and its position is shifted  
 508 from 275 nm to 480 nm. On the opposite, the position of fluorophore 3 is  
 509 unchanged, its global shape is almost correct but the respective magnitude  
 510 of its two peaks is largely modified. CDA-PARAFAC provides an impressive  
 511 correction of both deviations.

512 Actually, CDA-PARAFAC appears to be a critical improvement of U-PARAFAC  
 513 or even ACA-PARAFAC in the case of strong inner filter effects.

## 514 5 Conclusion

515 It is possible to correct inner filter effects by using simply a controlled dilution  
516 approach (CDA). This analytical solution is better than usual absorbance cor-  
517 rection and quicker and safer than strong dilution, under absorbance of 0.3.  
518 It has been demonstrated in this work the good ability of CDA and CDA-  
519 PARAFAC in the case of standard mixtures of two fluorophores and in the  
520 case of real DOM samples. In this study, CDA performed very well for solu-  
521 tion absorbances up to 1.83. Further investigation should be made outside this  
522 range. We conjecture that an other theoretical model of fluorescence measure-  
523 ment should be used for absorbance higher than 2.  
524 In respect to the PARAFAC decomposition, better results were obtained on  
525 the spectral loadings. We have also highlighted the limit of ACA for the cor-  
526 rection of strong filter effects. Consequently, we recommend the use of CDA  
527 for FEEM experiment and PARAFAC pretreatment to avoid error and misin-  
528 terpretation.

## 529 References

- 530 [1] B. Valeur, *Molecular Fluorescence. Principles and Applications*, Wiley-VCH,  
531 Weinheim, 2002.
- 532 [2] A. Hansch, D. Sauner, I. Hilger, J. Böttcher, A. Malich, O. Frey, R. Bräuer and  
533 W.A. Kaiser, *Academic Radiology*, 11 (11) (2004) 1229-1236.
- 534 [3] E. Sikorska, T. Gorecki, I.V. Khmelinskii, Sikorski and D. de Keukeleire, *Food*  
535 *chemistry*, 96 (4) (2006) 632-639.
- 536 [4] G.P., Coble, *Marine Chemistry*, 51 (1996), 325-346.
- 537 [5] C.A. Stedmon, S. Markager, *Limnology and Oceanography*, 50 (2) (2005).
- 538 [6] J.R., Lakowicz, *Principles of Fluorescence Spectroscopy*, Plenum Press, NY,  
539 1983.
- 540 [7] C.A. Parker, W.J. Barnes, *The analyst*, 82 (1957) 606-618.
- 541 [8] J.J. Mobed, S.L. Hemmingsen, J.L. Autry, and L.B. McGown, *Environmental*  
542 *Science and Technology*, 30 (10) (1996) 3061-3065.
- 543 [9] J.F. Holland, R.E. Teets, P.M. Kelly and A. Timnick, *Analytical Chemistry*, 49  
544 (6) (1977) 706-710.
- 545 [10] C. M. Yappert and J.D. Ingle, *Applied Spectroscopy*, 43 (5) (1989) 759-767.
- 546 [11] S.A. Tucker, V.L.Amszi, W.E. Acree, *Journal of Chemical Education*, 69 (1992).

- 547 [12] B.C. MacDonald, S.J.Lvin and H. Patterson, *Analytica Chimica Acta*, 338  
548 (1997) 155-162.
- 549 [13] A. Credi, and L. Prodi, *Spectrochimica Acta Part A*, 54 (1998) 159-170.
- 550 [14] J. Riesz, J. Gilmore, P. Meredith, *Spectrochimica Acta part A* (2004).
- 551 [15] T. Ohno, *Environmental Science and Technology*, 36 (4) (2002) 742-746.
- 552 [16] D.M. MacKnight, E.W Boyer, P.K. Westerhoff, P.T. Doran, T. Kulbe and D.T.  
553 Andersen, *Limnology and Oceanography*, 46 (2001) 38-48.
- 554 [17] F. Cuesta Sanchez, S. Rutan, M. Gil Garcia and D. Massart, *Chemometrics  
555 and Intelligent Laboratory Systems*, 36 (1997) 153-164.
- 556 [18] A. Garrido Frenich, D. Picon Zamora, J. Martinez Vidal and M. Martinez  
557 Galera, *Analytica Chimica Acta*, 449 (2001) 143-155.
- 558 [19] J. Boehme, P. Coble, R. Conmy and A. Stovall-Leonard, *Marine Chemistry*, 89  
559 (2004) 3-14.
- 560 [20] J.M. Andrade, M.P. Gomez-Carracedo, W. Krzanowsky and M. Kubista,  
561 *Chemometrics and Intelligent Laboratory Systems*, 72 (2004) 123-132.
- 562 [21] R.A. Harshman, *UCLA Working Papers in Phonetics*, 16 (1970) 1-84 (UMI  
563 Serials in Microform, No. 10,085).
- 564 [22] R. Bro, *Chemometrics and Intelligent Laboratory Systems*, 38 (1997) 149-171.
- 565 [23] C.A. Stedmon, S. Markager and R. Bro, *Marine Chemistry*, 82 (3-4) (2003)  
566 239-254.
- 567 [24] R.D. Holbrook, J.H. Yen and T.J. Grizzard, *Science of The Total Environment*,  
568 361 (1-3) (2006) 249-266.
- 569 [25] X. Luciani, S. Mounier, H.H.M. Paraquetti, R. Redon, Y. Lucas, A. Bois, L.D.  
570 Lacerda, M. Raynaud and M. Ripert, *Marine Environmental Research*, 65 (2)  
571 (2008) 148-157.
- 572 [26] D. Baunsgaard, Department of Dairy and Food Science, The Royal Veterinary  
573 and Agricultural University, (1999).
- 574 [27] G.G. Anderson, K.D. Brian and K.S. Booksh, *Chemometrics and Intelligent  
575 Laboratory Systems*, 49 (1999) 195-213.
- 576 [28] O. Divya and A.K. Mishra, *Analytica Chimica Acta*, 630 (2008) 47-56.
- 577 [29] P. John, I. Soutar, *Analytical Chemistry*, 48 (1976) 520-524.
- 578 [30] D. Patra and A.K. Mishra, *Analyst*, 125 (2000) 1383-1386
- 579 [31] R.A. Harshman, *UCLA Working Papers in Phonetics*, 22 (1972) 111-117 (UMI  
580 Serials in Microform, No. 10,085).
- 581 [32] N.D. Sidiropoulos and R. Bro, *Journal of Chemometrics*, 14 (3) (2000) 229-239.

- 582 [33] N.M. Faber, R. Bro and P.K. Hopke, *Chemometrics and Intelligent Laboratory*  
583 *Systems*, 65 (2003) 119-137.
- 584 [34] G. Tomasi, R. Bro, *Computational Statistics & Data Analysis*, 50 (7) (2006)  
585 1700-1734.
- 586 [35] P. Paatero, *Chemometrics and Intelligent Laboratory Systems*, 38 (2) (1997)  
587 223-242.
- 588 [36] R.A. Harshman and M.E. Lundy, *Computational Statistics & Data Analysis*,  
589 18 (1994) 39-72.
- 590 [37] C. Andersen, R. Bro, *Journal of Chemometrics*, 17 (2003) 200-215.
- 591 [38] R.G. Zepp, W.M. Sheldon and M.A. Moran, *Marine Chemistry*, 89 (2004) 15-36.
- 592 [39] C.A. Andersson and R. Bro, *Chemometrics and Intelligent Laboratory Systems*,  
593 52 (2002) 1-4.
- 594 [40] R. Bro and H.A.L. Kiers, *Journal of Chemometrics*, 17 (2003) 274-286.