Manuscript after review

Access to published version: https://www.sciencedirect.com/science/article/pii/S030881461530042X

FLUORESCENCE PROPERTIES OF FLAVONOID COMPOUNDS. QUANTIFICATION IN PAPRIKA SAMPLES BY USING SPECTROFLUORIMETRY COUPLED TO SECOND ORDER CHEMOMETRIC TOOLS

Olga Monago-Maraña, Isabel Durán-Merás, Teresa Galeano-Díaz*, Arsenio Muñoz de la Peña

Department of Analytical Chemistry and IACYS, University of Extremadura, Badajoz, 06006, Spain

* Corresponding author. E-mail address: tgaleano@unex.es

Running title: Fluorimetry coupled to second order chemometric tools to quantify flavonoids

Abstract

The influence of pH over fluorescence of flavonoid compounds was investigated and the highest fluorescence emission was obtained in basic medium. Selected conditions to improve this signal were: pH 9.5, 0.14 M Britton Robinson buffer and methanol between 5 - 10 %. The excitation - emission fluorescence matrices of a set of 36 samples of Spanish paprika were analyzed by means of parallel factor analysis (PARAFAC). Thus, the profiles of possible fluorescence components (PARAFAC loadings) were obtained. One of these profiles was identified by matching PARAFAC scores with LC analysis on the same samples. Two clusters of samples were obtained when score values were plotted against each other. Spectrofluorimetry coupled to second order multivariate calibration methods, as unfolded-partial least squares with residual bilinearization (U-PLS/RBL) and multidimensional-partial least-squares with residual bilinearization (N-PLS/RBL), was investigated to quantify quercetin and kaempferol in those samples. Good results were obtained for quercetin by this approach.

Keywords: flavonoids, paprika, fluorescence, parallel factor analysis, unfolded-partial least-squares with residual bilinearization, multidimensional-partial least-squares with residual bilinearization

Chemical compounds studied in this article: Quercetin (PubChem CID: 5280343); Myricetin (PubChem CID: 5281672); Kaempferol (PubChem CID: 5280863).

1. Introduction

Paprika is a red powder obtained by grinding the dried pepper pods of some varieties of *Capsicum annuum L*. This natural food product is commonly used as spice and natural colorant in cookery and to provide redness to meat products and commercial sauces (Palacios-Morillo, Marcos Jurado & Alcázar, 2014). It is a vegetable derived product which is particularly rich in organic microcomponents with antioxidant properties. Tocopherols, capsaicinoids, flavonoids and carotenes belong to these antioxidants present in paprika.

Flavonoids are a group of polyphenolic compounds distributed in medicinal plants, vegetables, fruit juices and beverages (tea, coffee, wines…) (Liu & Guo, 2006, Shaghaghi, Manzori & Joyban, 2008, Shaghaghi, Manzori, Afshar & Joyban, 2009). These compounds show high antioxidant and anticancer activities, which are determined by the presence of a number of hydroxyl groups at a certain positions, and a double bond at a C2-C3 position (Bae, Jayaprakasha, Jifon & Bhimanagouda, 2012). The positive effect of flavonoids against some diseases is attributed to the inhibition of specific enzymes, antioxidant activity, vascular protection, and to the anti-hepatotoxic, anti-allergic, anti-proliferative, anti-osteoporotic, and anti-inflammatory properties. Therefore, these compounds are potent regulators for cell cycle progression, which may be involved in the prevention of carcinogenesis (Ramesová, Sokdová, Degano, Bulícková, Zabka & Gal, 2012).

Flavonoid glycosides (flavonoids bound to various sugars) are found in paprika and peppers. Free flavonoid aglycones can be produced from these flavonoid glycosides as a result of hydrolysis of the glycosidic bond by enzymes or acidic conditions (Bae et al., 2012, Jeffery, Parker & Smith, 2008).

Regarding to the determination of these compounds in foods, separative techniques have been widely used. In this sense, flavonoids have been determined in tea, onions, wines, peppers, etc., by using liquid chromatography (LC) coupled to UV-Visible detection, diode-array-detection (DAD), mass spectrometry (MS) or magnetic resonance (NMR), or by using capillary electrophoresis coupled to UV detection. (Delgado, Tomás-Barberán, Talou & Gaset, 1994,

Ehala & Veher, 2005, Valls, Millán, Martí, Borrás & Arola, 2009, Rijke, Out, Niessen, Ariese, Gooijer & Brikman, 2006, Molnár-Perl & Füzfai, 2005, Careri, Bianchi & Corradini, 2002, Bae et al., 2012).

However, spectroscopic techniques are not frequently used for this purpose. Shanghaghi et al. (2008) quantified the total content of flavonoids in foods by a fluorescence method based on terbium complexation. On the other hand, Perucka and Materska (2003) quantified the total content of flavonoids by using spectrophotometry, and, Zaki, Hakmaoui, Ouatmane, Hasib and Fernández - Trujillo (2013) determined spectrophotometrically the total content of flavonoids, by means of the formation of flavonoid-aluminium complexes.

Fluorescence is a sensitive and selective analytical technique. In the last years, the application of fluorescence for analysis of complex samples such as foods has increased due to the possibility of combining the technique with chemometric tools. In addition, second-order algorithms present an advantage, which is the ability to get accurate concentration estimates of analytes of interest, even in the presence of uncalibrated interfering components, which should allow for an improvement in predictive ability (Escándar, Goicoechea, Muñoz de la Peña & Olivieri, 2014, Muñoz de la Peña, Olivieri, Escándar & Goicoechea, 2015).

In this sense, fluorescence coupled to PARAllel FACtor analysis (PARAFAC) has been used for the characterization and classification of wines (Airado-Rodríguez, Galeano-Díaz & Durán-Merás, 2009) and honey (Lenhardt, Bro, Zekovic, Dramicanin & Dramicanin, 2015) samples. Similarly, synchronous fluorescence and multivariate classification analysis have been recently used to determine Sudan I, a colorant employed in industrial applications, in culinary spices (Di Anibal, Rodríguez & Albertengo, 2015). In the case of analysis of flavonoids in a complex matrix, such as is the paprika, studies using spectroscopic techniques were not found.

On the other hand, other strategy in working with second-order data is to rearrange them in vectors and then apply a first-order algorithm such as unfolded partial-squares (U-PLS) or the multi-dimensional variant (N-PLS). Unfolding the matrix calibration data leads to the

possibility of applying classical partial least-squares (PLS), a popular regression technique in the framework of first-order calibration. The achievement of the second-order advantage is left to a post-calibration procedure called residual bilinearization (RBL), which processes the test samples in the original matrix form, efficiently separating the contribution of the potential interferents from those of the calibrated analytes. The resulting U-PLS/RBL algorithm shows a great flexibility, and is able to cope with some data sets deviating from trilinearity. (Olivieri, Escándar, Goicoechea & Muñoz de la Peña, 2015).

From this background, taking into account the importance of avoiding fraud in this kind of samples which are recognized under a Protected Designation of Origin (PDO), the aims of this study were the following: exploring the possibilities of the fluorescence properties of flavonoid compounds in order to their analysis, trying to differentiate paprika samples according to their origin, with the base of their total content of flavonoids and, finally, developing an alternative method for quantifying a mixture of flavonoids in paprika samples using spectrofluorimetry coupled to second order algorithms (PARAFAC, unfolded-partial least-squares with residual bilinearization (U-PLS/RBL) and multidimensional partial least-squares with residual bilinearization (N-PLS/RBL), utilized for the first time for these compounds in this food matrix.

2. Materials and methods

2.1. Chemical reagents and samples

Apigenin, luteolin, myricetin and kaempferol standards were purchased from Extrasynthese (Genay Cedex, France). LC-grade methanol solvent and quercetin standard were obtained from Sigma (Sigma-Aldrich Química, S.A., Madrid).

Britton-Robinson (BR) buffers of different pH were prepared from a 0.04 M acetic acid, 0.04 M phosphoric acid and 0.04 M boric acid solution in 100 mL calibrated flasks, and the corresponding volume of 0.02 M NaOH to obtain the appropriate pH.

105 Samples of paprika belonging to different origins: Spanish Protected Designation of Origin
106 (PDO) "*Pimentón de La Vera*" and other different producers, which were obtained from
107 Regulatory Council of the Denomination of Origin "*Pimentón de La Vera*" and from market,
108 respectively.

109 **2.2. Instrumentation and software**

110 Fluorescence measurements were performed on a Cary Eclipse VARIAN spectrofluorometer
111 equipped with two Czerny-Turner monochromators, a xenon light source and a photomultiplier
112 tube as detector. A 1.0 cm quartz cell was used. The Cary Eclipse software was used for data
113 acquisition. To obtain fluorescence excitation-emission matrices (EEMs), excitation
114 wavelengths were increased from 400 to 470 nm at 5 nm steps; for each excitation wavelength,
115 the emission spectrum was obtained in the range 480 - 600 nm at 2 nm. The instrumental
116 parameters used were as follow: 650 V and slit widths of 10 nm. Moreover, emission spectra
117 were smoothed using the Savitzky-Golay method (5 experimental points).

118 The pH of the solutions was measured with a Crison MicopH 2001 meter (Barcelona, Spain),
119 equipped with a combined glass/saturated calomel electrode.

120 The software package The Unscrambler® v6. 11 (CAMO A/S Olav Tryggvasonsgt, N-7011,
121 Trondheim, Norway) was used for the experimental designs.

122 Analysis of data were done using MatLab R2008a (MATLAB Version 7.6, The Marhworks,
123 Natick, Massachusetts, 2010), the MVC2 routine developed by Oliveri (Oliveri, Wu & Yu,
124 2009) and the PLS toolbox routine (Eigenvector Research Inc., Wenatchee, WA). ACOC
125 program was used for statistical analysis (Espinosa Mansilla, Muñoz de la Peña & González
126 Gómez, 2005).

127 **2.3. Samples treatment**

128 The analytes were extracted from 0.5 g of paprika sample with 20 mL of MeOH for 30 min in
129 an ultrasonic bath. The extract solution was centrifuged and evaporated to dryness. The residue
130 was suspended in water and loaded on a C18 cartridge (Solid Phase Extraction). The cartridge

was successively washed with 20 mL of water and 20 mL of 10% aqueous methanol to eliminate a part of paprika matrix. The analytes were eluted with 2.5 mL of MeOH 85%. Finally, they were in a final volume of 3.0 mL.

An aliquot of the extract eluted from the cartridge was subjected to hydrolysis in a water bath at 85 ºC for 45 min with a final concentration of HCl in the flask of 2.5 M, to obtain aglycones of flavonoids according to a previously optimized procedure (data sent to publish). Excitation-emission fluorescence matrices were obtained for samples prepared as follow: 0.4 mL of the hydrolyzed was diluted with a Britton –Robinson buffer solution of pH 12.9 to obtain a pH around 9.5 – 10, in a final volume of 3.0 mL. In the Fig. S1 a scheme of the experimental procedure for the sample treatment is shown.

2.4. Data modelling for PARAFAC analysis

Initially, with the aim of evaluating the capabilities of EEMs to distinguish between samples of different origin, a PARAFAC model was constructed using the EEMs of a set of 17 samples of *"Pimentón de la Vera"* paprika samples, and 19 of paprika samples from other Spanish producers. A pretreatment of data set to remove the Rayleigh signals in all the EEMs used for PARAFAC analysis was performed according to Airado-Rodríguez et al. (2009). Subsequently, to model the set of fluorescence data by PARAFAC, the EMMs of the 36 samples were arranged in a three-dimensional structure of size 36 x 60 x 15 (samples x number of emission wavelengths x number of excitation wavelengths). This array was decomposed by PARAFAC using different number of components. In all cases, non-negative constraints for the resolved profiles for all modes were applied with the purpose to obtain a realistic solution, because concentrations and spectral values are positive.

2.5. Calibration and Test Sets

To assess the ability of the different second-order multivariate analysis tools, in the determination of a mixture of the two main fluorescent flavonoids of paprika, quercetin and kaempferol, an 11-standards set was built for calibration with the PARAFAC, U-PLS/RBL and N-PLS/RBL models. The analyte concentrations, from 0.05 to 0.4 $\mu g \cdot mL^{-1}$, corresponded to a

central composite design. Samples were prepared in 3.0 mL calibrated flasks in presence of 7 % of methanol and diluting to the mark with BR buffer solution (0.16 M, final pH = 9.5). Moreover, a set of six test samples was prepared with analyte concentrations different from those employed for calibration but within their corresponding calibration ranges. EEMs were measured in a range of 400 – 470 nm for excitation wavelengths and in a range of 480 – 600 nm for emission wavelengths.

Another 16-sample set was built for calibration with the PARAFAC, U-PLS/RBL AND N-PLS/RBL models, in order to be used for analysis of real paprika samples. The analyte concentration used for calibration corresponded to a fractional factorial design, and concentration levels were selected according their contents in paprika samples, ranging from 0.04 to 0.4 $mg \cdot L^{-1}$ of quercetin and from 0.0067 to 0.03 $mg \cdot L^{-1}$ of kaempferol. This design provided a total of seven standards. Moreover, nine additional standards were added containing only quercetin or kaempferol in order to provide more information of the pure analytes.

Also, an additional set of six test samples was built with analyte concentrations different from those employed for calibration, but within their corresponding calibration ranges.

3. Results and discussion

3.1. Fluorescence behaviour studies

Firstly, influence of pH over fluorescent behaviour of quercetin was examined. Emission spectra were recorded, in the range 480 – 600 nm, using excitation wavelength of 440 nm, at different pH values (Fig. 1A). It can be observed that basic medium improves fluorescence signal of quercetin, however the stability of this signal decreases in this medium, which is in accordance to data described in the literature (Ramesová et. al., 2012). A kinetic study about fluorescence behaviour depending on pH value was made. It was observed that at pH higher than 10.8 the fluorescence signal quickly decreases with time (Fig. 1B). An experiment with a higher percentage of methanol was performed and stability was better, but the signal was lower.

According to these results, it was decided to select a working pH of 9.5, in order to avoid instability, and optimizing other experimental conditions. On the other hand, instrumental parameters of the spectrofluorimeter were optimized to improve fluorescence signal and those indicated in the section 2.2. were selected.

In order to study the influence of the percentage of methanol, this was varied in the range 5 – 80%. It was observed that low percentages of methanol offered higher fluorescence signals and, in consequence, we decided to work in the range of 5 – 10 %, according to the necessary conditions of the real samples analyzed.

Another experimental optimized condition was the volume of buffer solution. By using a Britton Robinson buffer solution of pH 10 (0.16 M), the volume of it was varied in a range of 0.6 – 2.6 mL in a final volume of 3 mL (buffer concentration in the range 0.03 – 0.14 M). It was tested that the signal increased when concentration of buffer did it. For this reason, we decided to adjust percentage of methanol to 7%, diluting to the mark with this buffer solution. The fluorescence spectrum of a blank solution was registered to check that the signal was due to quercetin.

In these conditions, an EEM of quercetin was obtained in the excitation range 230 – 480 nm and in the emission range of 250 – 600 nm. The better signals were obtained for an excitation wavelength range of 400 – 480 nm and an emission wavelength range of 500 – 600 nm.

3.2. Fluorescence behaviour of other flavonoids

EEMs of other flavonoids referenced as paprika components, kaempferol, myricetin, luteolin and apigenin were registered in the same conditions. However, only three out of the five compounds studied showed fluorescence (quercetin, kaempferol and myricetin), which are called flavonols. The EEMs of these compounds are shown in the Figure 2. Moreover, the stability of their fluorescence signals was examined and it was found that they were stable at least over ten minutes.

This behaviour can be explained taking into account that it is known that the H-bond between C(4)=O···HO-C(5) (present in the five compounds studied) may favour the non-radiative deactivation, while that between C(4)=O···HO-C(3) (in quercetin, myricetin and kaempferol) may permit the excited-state proton transfer process, given rise to a tautomeric equilibrium responsible of two different bands in the fluorescence emission spectra (Cao et al., 2014). This way, use of surfactants for solubilization of both C(4)=O and OH-C(3) has been employed to improve fluorescence signal of these compounds (Cao et al., 2014; Liu et al., 2006). In this case, surfactants have not been used, however the fluorescence signal has notably increased with the pH.

In Figure 3A, it can be observed the deprotonation procedure of the three fluorescent flavonoids, according to the literature (Álvarez-Diduk, Ramírez-Silva, Galano & Merkoci, 2013), as well as the range of values referenced for pKa for each deprotonation. Regarding to these pKa values, it can be said that at a pH value around 9.5, second and third deprotonation are occurring, in the case of myricetin and quercetin and, first and second ones, are occurring in the case of kaempferol.

It may be thought that the fluorescence is due to the forms B or C, because these can be present for the three compounds at this pH value. Nevertheless, a notable increasing of the fluorescence signal was observed when pH was higher than 9, so it can be said that the fluorescence signal is due to the second deprotonation. According the literature, the first deprotonation might be related with the presence of the ground-state of the enolic tautomer in the medium (Figure 3B) (Smith & Markham, 1998). It can be deducted that the second deprotonation of these compounds stimulates this form and, as a result, the fluorescence signal. As reflected in Figure 3B, this may be due to a second resonating part that is formed in the molecule.

On the other hand, differences in fluorescence intensity observed for the three compounds could be explained considering various factors. One of them may be understood in relation to the presence of -OH groups in the ring B, because these groups would disfavour the resonant form, which is responsible of the fluorescence signal. Another one, it might be deducted according to

237 the proportion of the molecule responsible of the fluorescence signal, which, taking into account
238 the pKa values found in the literature (Álvarez-Diduk et al., 2013), could be: kaempferol >
239 quercetin > myricetin. Moreover, the pKa value corresponding to the third deprotonation
240 decreases in the order myricetin, quercetin and kaempferol. At our working pH, this
241 deprotonation is occurring and, in consequence, myricetin would be deprotonated in a great
242 extension. Lastly, the influence of oxidation of all of these structures in basic medium,
243 described in literature (Ramesová et. al., 2012), could also affect the amount of the fluorescent
244 one in different grades.

245 **3.3. PARAFAC analysis for samples differentiation**

246 The possibility of using EEMs to distinguish between samples belonging to different categories
247 has been examined. With this purpose, multivariate data analysis was performed by using
248 PARAFAC. EEMs were obtained for a set of 36 samples, 17 of which are paprika samples from
249 the PDO "Pimentón de la Vera" (Cáceres, Spain), and other 19 are samples from different
250 producers acquired in the market. In the first place, we chose the appropriate number of
251 components for constructing the PARAFAC model. The criterion that we used was the core
252 consistency diagnostic (Andersen & Bro, 2003). Core consistency percentages of 100, 98.11,
253 96.6 and -205 % were obtained for one, two, three and four component models, respectively. It
254 was clear that three components were the optimal in the present case for constructing the model.
255 The 3D structures of the first two PARAFAC components obtained from this PARAFAC model
256 are shown in the Figure 4A.

257 According to the loading shapes, it might say that the 3D loading corresponding to the second
258 component is similar to the EEMs of flavonoid compounds (Figure 2). For this reason, we
259 studied the possible correlations between PARAFAC scores and the flavonoid concentrations of
260 these samples, quantified by a LC method previously developed (data send to publish).

261 For the first component, it was not found a good correlation between the score values and the
262 quercetin, kaempferol or luteolin concentration. It is concluded that this signal corresponds to
263 other fluorescence compounds present in this matrix.

For the second component, correlation was not good either between score values and luteolin concentration. This result is logical given that luteolin does not present fluorescence signal in the selected experimental conditions. For kaempferol concentration, poor correlation was found. However, for the second component, high correlations were found between the scores and quercetin concentrations (R = 0.9180) (Figure 4B). In addition, good correlations were found among the scores and quercetin + kaempferol concentrations (R = 0.9180). This result is due to the fact that kaempferol concentrations are very low and they do not affect the correlation.

The score values corresponding to each PARAFAC component were plotted against each other in order to study possible systematic information contained in fluorescence data, with respect to the variable origin of the sample. Scores corresponding to the first and second PARAFAC components reveal two clusters of the paprika samples, according to their origin (Figure 5). PDO samples have higher values of scores for the first component; however, there are not differences between score values for the second and third components.

3.4. Resolution and quantification of a mixture of flavonoids by using PARAFAC and U-PLS/RBL and N-PLS/RBL models

Regarding to previous performed studies, the resolution of a mixture of different flavonoids by using chemometric tools was investigated. This study was focused in quercetin and kaempferol, because luteolin, which is also present in paprika, was not fluorescent, and myricetin was not present in Spanish paprika samples. Previously, it was proven that the quercetin (fluorescent major component in paprika sample) could not be determined by using external or addition standard methodologies, even when a cleaning stage of the samples with a C18 cartridge was performed. Hence, obtained results by these first-order methodologies were not in accordance with those obtained by LC. For these reason, it was decided to employ second-order algorithms to model the present interferences.

3.4.1. Selection of the number of factors and validation of the model with synthetic samples

290 In order to optimize the second-order multivariate models mentioned in the section 2.5, the
291 selection of the number of factors was performed. Pursuing this goal, a set of 11 calibration
292 samples was employed, which containing a mixture of quercetin and kaempferol, in a range of
293 0.05 - 0.4 mg· L^{-1}. For U-PLS and N-PLS, cross-validation and the Haaland and Thomas
294 criterion (Haaland & Thomas, 1988) were employed to choose the optimum number of factors.
295 The number of factors corresponding to the model given a PRESS value statistically no
296 different to the minimum PRESS value (F-ratio probability falling below 0.75) was selected as
297 the optimum. In this case, three factors were found for both methods, and for each component of
298 the mixture.

299 With the aim of validating the proposed chemometric methods, a set of 6 test samples
300 containing a mixture of quercetin and kaempferol, in the same range of concentrations that the
301 calibration samples, was analyzed. In the table S1, the results of the validation with U-PLS/RBL
302 and N-PLS/RBL methods are shown. It can be observed that the results of the analysis of the
303 synthetic samples corresponding to the validation set are satisfactory for the two methods, with
304 mean recovery values, in percentage, ranging from 89 to 120 %.

305 In PARAFAC, to select the optimal number of factors, the set of test samples was used. When
306 the variation of the core consistency was plotted versus the trial number of components, the core
307 value drops below zero when the number of factors used was higher than one, which is
308 conclusive that it could not difference both analytes.

309 Because it was not possible to quantify both analytes separately, we tried to predict the factor
310 obtained as the total content of flavonoid compounds. In the table S2, the results of PARAFAC
311 method are shown. It can be observed that the results of the analysis of the validation set are
312 enough good for the total content of flavonoids, in the synthetic samples analysed, with mean
313 recovery values, in percentage, ranging from 69 to 119 %.

314 Pursuing the goal of using these methods for real paprika samples, another calibration set was
315 constructed corresponding to real concentrations in the samples, quercetin in the range of 0.04 -

316 0.4 $mg \cdot L^{-1}$ and kaempferol in the range 0.0067 - 0.03 $mg \cdot L^{-1}$. In the case of the selection of the
317 optimum number of factors, the results were the same. The obtained recovery values of U-
318 PLS/RBL and N-PLS/RBL obtained, for tests samples (7 – 12), are shown in the Table S1; all
319 are between 74 – 120 %. Therefore, the results show that U-PLS/RBL offered better results in
320 the case of kaempferol. In the case of PARAFAC, the results are presented as total content of
321 flavonoids, in the synthetic samples analysed, with recovery values, in percentage, ranging from
322 76 – 126 % (data shown in the Table S2).

323 Statistical results were also calculated. In the case of U-PLS, for quercetin, the root mean square
324 error of prediction (RMSEP) was 0.026 $mg \cdot L^{-1}$ and the relative error of prediction (REP) was 19
325 %, and, for kaempferol the RMSEP was 0.003 $mg \cdot L^{-1}$ and the REP was 17%. Regarding to N-
326 PLS, the RMSEP were 0.026 and 0.003 $mg \cdot L^{-1}$ for quercetin and kaempferol, respectively, and
327 the REPs were 18% for both analytes.

328 **3.5. Analysis of real paprika samples**
329
330 A group of real samples was analysed by using the three multivariate calibration models
331 described above. The samples were divided in two groups: belonging to PDO *"Pimentón de La*
332 *Vera"*, or not belonging. When applying U-PLS/RBL and N-PLS/RBL to the paprika samples,
333 it was necessary to assess the number of unexpected components to be employed in the RBL
334 procedure (Jiménez Girón, Durán-Merás, Muñoz de la Peña, Espinosa Mansilla & Cañada
335 Cañada, 2008; Olivieri, Escandar & Muñoz de la Peña, 2011). The number of unexpected
336 components was different in the case of quercetin or kaempferol. The results were a single new
337 factor, besides those required for calibration, for quercetin, and two new factors for kaempferol.
338 The new factors modelled the matrix signal of paprika.

339 Concentrations of quercetin (between 44 – 200 $mg \cdot kg^{-1}$ for PDO samples and between 96 – 560
340 $mg \cdot kg^{-1}$ for no PDO samples) and kaempferol (between 0.24 – 7 $mg \cdot kg^{-1}$ for PDO samples and
341 between 0.26 – 11 $mg \cdot kg^{-1}$ for no PDO samples) obtained by using U-PLS/RBL and N-
342 PLS/RBL were correlated with concentrations obtained in an LC method (data send to publish),
343 which were good in the case of quercetin and, in the case of kaempferol, only 10 out of 36

samples were predicted correctly. Interestingly, the results were better in the case of the samples not belonging to PDO, which can be due to two reasons: the first one is that kaempferol is at lower concentrations than quercetin in the analysed samples and, the second one, is that matrix interferences are more important in the PDO samples, which is in accordance with the score values obtained for the first component in PARAFAC analysis, which are higher for the PDO samples than for the no PDO samples (Figure 5). In the Figure 6, it can be observed the correlation between quercetin concentrations predicted by second order algorithms and the quercetin concentrations determined by means of LC analysis.

In real samples, statistical results were obtained too. These results were better in the case of quercetin. In the first place, in the case of U-PLS/RBL, for PDO samples, RMSEP was 34 $mg \cdot kg^{-1}$ and REP was 25% and, for no PDO samples, RMSEP was 36 $mg \cdot kg^{-1}$ and REP was 17%. In the second place, in the case of N-PLS/RBL, for PDO samples, RMSEP was 34 $mg \cdot kg^{-1}$ and REP was 24% and, for no PDO samples, RMSEP was 36 $mg \cdot kg^{-1}$ and REP was 18%.

For PARAFAC analysis, RMSEP were 3.7 and 25 $mg \cdot kg^{-1}$ for PDO samples and no PDO samples, respectively, and REPs were 5.4 and 25 %, respectively. The results were better for PDO samples than for no PDO samples.

3.6. Quality assurance/quality control (QA/QC)

The quality control parameters commonly used to check the accuracy of an analysis for different compounds in food matrices are the calculated recoveries, corresponding to the additions of different fortifications of a standard.

On this way, the extraction procedure was validated by means of two procedures. One of them consisted of performing a second extraction of the remnant analyte in the sample, and the recoveries obtained were about 85% for quercetin and 80% for kaempferol, for the first extraction. On the other hand, the samples were fortified with known concentrations of quercetin and kaempferol and the obtained concentration were compared with the added. Recoveries results were very similar. The results of concentration data in both cases were

obtained by a HPLC method (results sent to publish). In addition, the other experimental steps (cleanup step and acid hydrolysis) were checked and the results showed recoveries round 100 %.

According to all this and the equation (1), the corrected concentrations founded in the samples should be, in the case of quercetin, between 52 – 235 $mg \cdot kg^{-1}$ for PDO samples and between 113 – 660 $mg \cdot kg^{-1}$ for no PDO samples, and, in the case of kaempferol should be between 0.3 – 9 $mg \cdot kg^{-1}$ for PDO samples and between 0.3 – 14 $mg \cdot kg^{-1}$ for no PDO samples.

$$C_c = \frac{C_f \, x \, 100}{\%R} \quad (1)$$

In Eq. (1) C_c is the corrected concentration of analyte, C_f is the founded concentration, D.F. is the dilution factor which would be applied to the samples, including all dilutions and the weight of sample and %R is the percent recovery of analytes in the sample, calculated as described above.

With this in mind, the limits of detection (LOD_c) and the limits of quantification (LOQ_c) were calculated from the LOD and the LOQ that the U-PLS and N-PLS method offered for each standard. Recoveries obtained in each experimental step were considered and the equations 1 and 2 were finally employed for this purpose.

$$LOD_c = \frac{D.F.x \; LOD \; (U-PLS \; or \; N-PLS) \; x \; 100}{\%R} \quad (2)$$

$$LOQ_c = \frac{D.F.x \; LOQ \; (U-PLS \; or \; N-PLS) \; x \; 100}{\%R} \quad (3)$$

In Eq. (2) and (3), LOD and LOQ are the limits of detection and quantification of the method, respectively, calculated with test samples, D.F. is the dilution factor which would be applied to the samples, including all dilutions and the weight of sample and %R is the percent recovery of analytes in the sample, calculated as described above.

393 According to all of this, the LODc obtained were in the range of 1 – 3 $mg \cdot kg^{-1}$ and 4 – 9 $mg \cdot kg^{-1}$
394 for quercetin by using U-PLS/RBL and N-PLS/RBL, respectively. The LOQc were in the range
395 of 6 – 8 $mg \cdot kg^{-1}$ and 20 – 24 $mg \cdot kg^{-1}$ for quercetin by using U-PLS/RBL and N-PLS/RBL,
396 respectively. In the case of kaempferol, the LODc were in the range of 1 – 2 $mg \cdot kg^{-1}$ and 4 – 5
397 $mg \cdot kg^{-1}$ for U-PLS/RBL and N-PLS/RBL, respectively and the LOQc were in the range of 4 – 5
398 $mg \cdot kg^{-1}$ and 12 – 14 $mg \cdot kg^{-1}$ for U-PLS/RBL and N-PLS/RBL, respectively.

399 However, we consider a more realistic estimation of LOQ the selection of the lowest
400 concentration for which good correlation between results of this method and HPLC results is
401 obtained when analyzing real sample. This way, the LOQ would be about 50 mg/kg and 10
402 mg/kg for quercetin and kaempferol, respectively.

403

404 **4. Conclusions**

405 The fluorescence properties of flavonoid compounds have been investigated and optimized, and
406 good results of intensity and stability for quercetin, myricetin and kaempferol compounds were
407 obtained. In addition, these properties have been intensified in basic medium for the first time.

408 A PARAFAC analysis has been performed with a set of 36 samples and they had been grouped,
409 according their belonging or not to a Spanish PDO, obtaining a clustering of them when the
410 score values of the first component against the score values of the second one were plotted. It
411 must be emphasized that the relation between the second loading and the flavonol contents was
412 proved, since there was a good correlation between scores of this loading and the concentration
413 of these compounds obtained by HPLC.

414 The usefulness of chemometric tools to identify and quantify a mixture of the selected
415 flavonoids in a complex matrix, as paprika samples, has been investigated. PARAFAC offers
416 the possibility of quantifying quercetin plus kaempferol, together, in presence of matrix
417 interferences. So, this method can be used to quantify flavonol compounds in paprika samples.
418 U-PLS/RBL and N-PLS/RBL allow differentiating quercetin and kaempferol in synthetic

samples, and quantifying them in paprika samples and, particularly, quercetin, because of its higher abundance in these samples.

Acknowledgements

The authors are grateful to the Ministerio de Economía y Competitividad of Spain (Project CTQ2014-52309-P) and the Gobierno de Extremadura (GR15-Research Group FQM003), both co-financed by the European FEDER funds, for financially supporting this work. Olga Monago Maraña is grateful to the Ministerio de Educación, Cultura y Deporte of Spain for a FPU grant (Resolución de 18 de noviembre de 2013, de la Secretaría de Estado de Educación, Formación Profesional y Universidades, BOE nº 279, de 21/11/13, reference number FPU13/02249).

References

Airado-Rodríguez, D., Galeano-Díaz, T. & Durán-Merás, I. (2009). Usefulness of fluorescence excitation-emission matrices in combination with PARAFAC, as fingerprint of red wines. Journal of agricultural and food chemistry, 57, 1711-1720.

Álvarez-Diduk, R., Ramírez-Silva, M.T., Galano, A. & Merkoci A. (2013). Deprotonation mechanism and acidity constants in aqueous solution of flavonol: a combined experimental and theoretical study. Journal of Physical Chemistry B, 117, 12347 – 12359.

Andersen, C.M. & Bro, R. (2003). Practical aspects of PARAFAC modelling of excitation-emission data. Journal of Chemometrics, 17, 200 - 215.

Bae, H., Jayaprakasha, G.K., Jifon, J. & Bhimanagouda, S.P. (2012). Extraction efficiency and validation of an HPLC method for flavonoid analysis in peppers. Food Chemistry, 130, 751 – 758.

Cao, W., Hu, S.S., Li, X-Y, Pang, X-Q., Cao, J., Ye, L-H., Dai, H-B., Liu, X-J., Da, J-H. & Chu, C. (2014). Highly sensitive analysis of flavonoids by zwitterionic microemulsion electrokinetic chromatography coupled with light-emitting diode-induced fluorescence detection. Journal of Chromatography A, 1358, 277 – 284.

Careri, M., Bianchi, F. & Corradini, C. (2002). Recent advances in the application of mass spectrometry in food-related analysis. Journal of Chromatography A, 970, 3-64.

Delgado, C., Tomás-Barberán, F.A., Talou, T. & Gaset, A. (1994). Capillary electrophoresis as an alternative to HPLC for determination of honey flavonoids. Chromatographia, 38, 1/2, 72 – 78.

Di Anibal, C.V., Rodríguez, M.S. & Albertengo L. (2015). Synchronous fluorescence and multivariate classification analysis as a screening tool for determining Sudan I dye in culinary spices. Food Control, 56, 18 – 23.

Ehala, S., Vaher, M. & Kaljurand, M. (2005). Characterization of phenolic profiles of Northern European Berries by capillary electrophoresis and determination of their antioxidant activity. Journal of Agricultural and Food Chemistry, 53, 6484-6490.

Escandar, G.M., Goicoechea, H.C., Muñoz de la Peña, A. & Olivieri A.C. (2014). Second- and higher-order data generation and calibration: A tutorial. Analytical Chimica Acta, 806, 8 - 26.

Espinosa Mansilla, A., Muñoz de la Peña, A. & González Gómez, D. (2005). Using univariate linear regression calibration software in the MATLAB environment. Application to chemistry laboratory practices. Chemical Education, 10, 337 – 345.

Haaland D.M. & Thomas E.V. (1988). Partial Least-Squares Methods for Spectral Analyses. 1. Relation to Other Quantitative Calibration Methods and the Extraction of Qualitative Information. Analytical Chemistry, 60, 1193 – 1202.

Haaland D.M. & Thomas E.V. (1988). Partial Least-Squares Methods for Spectral Analyses. 2. Application to Simulated and Glass Spectral Data. Analytical Chemistry, 60, 1202 – 1208.

Jeffery, D.W., Parker, M. & Smith, P.A. (2008). Flavonol composition of Australian red and white wines determined by high-performance liquid chromatography. Australian Journal of Grape and Wine Research, 14, 153 – 161.

473 Jiménez Girón, A., Durán-Meras, I., Muñoz de la Peña, A., Espinosa Mansilla, A., Cañada
474 Cañada F. & Olivieri A.C. (2008). Photoinduced fluorimetric determination of folic acid
475 and 5-methyltetrahydrofolic acid in serum using the kinetic evolution of the emission
476 spectra accomplished with multivariate second-order methods. Analytical Bioanalytical
477 Chemistry, 391, 827 – 835.

478 Lenhardt, L., Bro, R., Zekovic, I., Dramicanin, T. & Dramicanin, M.D. (2015). Fluorescence
479 spectroscopy coupled with PARAFAC and PLAS DA for characterization and
480 classification of honey. Food Chemistry, 175, 284 – 291.

481 Liu, W. & Guo, R. (2006). Interaction of flavonoid, quercetin with organized molecular
482 assemblies of noionic surfactant. Colloid Surface A, 274, 192 – 199.

483 Molnár-Perl, I. & Füzfai, Zs. (2005). Chromatographic, capillary electrophoretic and capillary
484 electrochromatographic techniques in the analysis of flavonoids. Journal of
485 Chromatography A, 1073, 201-227.

486 Muñoz de la Peña, A., Olivieri, A.C., Escandar, G.M. & Goicoechea, H.C. (2015).
487 Fundamentals and analytical applications of multi-way calibration. Elsevier Editorial
488 (Chapter 7 and 8)

489 Oliveri, A.C., Wu, H-L & Yu, R-Q. (2009). MVC2: a MATLAB graphical interface toolbox for
490 second-order multivariate calibration. Chemometrics and intelligent laboratory systems,
491 96, 246-251.

492 Olivieri, A.C., Escandar, G.M. & Muñoz de la Peña, A. (2011). Second-and higher-order
493 multivariate calibration methods applied to non-linear data. Advantages and limitations
494 of the different algorithms. Trends in Analytical Chemistry, 30, 607 – 617.

495 Olivieri, A.C., Escandar, G.M., Goicoechea, H. C. & Muñoz de la Peña, A. (2015). Unfolded
496 and multiway partial least-squares with residual multilinearization: fundamentals. In A.
497 Muñoz de la Peña, H.C. Goicoechea, G.M. Escandar & A.C. Olivieri (Eds.),
498 Fundamentals and Analytical Applications of Multiway Calibration. (pp 347 – 364).
499 Elsevier editorial.

500 Palacios-Morillo, A., Marcos Jurado, J., Alcázar, A. & De Pablos, F. (2014). Geographical
501 characterization of Spanish PDO paprika by multivariate analysis of multielemental
502 content. Talanta, 128, 15-22.

503 Perucka, I. & Materska, M. (2003). Antioxidant activity and content of capsaicinoids isolated
504 from paprika fruits. Polish journal of food and nutrition sciences, 12/53, 15 – 18.

505 Ramesová, S., Sokolová, R., Degano, I., Bulícková, J., Zabka, J. & Gál, M. (2012). On the
506 stability of the bioactive flavonoids quercetin and luteolin under oxygen-free conditions.
507 Analytical and Bioanalytical Chemistry, 402, 975 – 982.

508 Rijke, E., Out, P., Niessen, W.M.A., Ariese, F., Gooijer, C. & Brinkman, U.A.Th. (2006).
509 Analytical separation and detection methods for flavonoids. Journal of Chromatography
510 A, 1112, 31 – 63.

511 Shaghaghi, M., Manzoori, J.L & Joyban, A. (2008). Determination of total phenols in tea
512 influsions, tomato and apple juice by terbium sensitized fluorescence as an alternative
513 approach to the Folin- Ciocalteu spectrophotometric method. Food Chemistry, 108, 695
514 – 701.

515 Shaghaghi, M, Manzoori, J.L, Afshar, D.J. & Jouyban A. (2009). Determination of flavonoids
516 in pharmaceutical preparations using Terbium sensitized fluorescence method. DARU,
517 17, 264 -268.

518 Smith, G.J. & Markham, K.R. (1998). Tautomerism of flavonol glucosides: relevance to plant
519 UV protection and flower colour. Journal of photochemistry and photobiology A:
520 Chemistry, 118, 99 – 105.

521 Valls J., Millán S., Martí, M.P., Borrás, E. & Arola, L. (2009). Advanced separation methods of
522 food anthocyanins, isoflavones and flavanols. Journal of Chromatography A, 1216,
523 7143 – 7172.

524 Zaki, N., Hakmaoui A., Ouatmane A., Hasib A & Fernández-Trujillo J.P. (2013). Bioactive
525 components and antioxidant activity of Moroccan Paprika (Capsicum annuum L.) at
526 different period of harvesting and processing. Journal of Biology, Agriculture and
527 Healthcare. 3, 1-8.

528

529

Figure captions

Figure 1. Emission spectra of quercetin (λ_{exc} = 440 nm) at different pH values (A). Kinetic curves of quercetin fluorescence at different pH values (B).

Figure 2. Excitation-emission matrices (EEMs) of flavonoid compounds (quercetin and myricetin 5 $\cdot 10^{-6}$ M and kaempferol 3.5 $\cdot 10^{-6}$ M) in basic medium (pH = 9.5).

Figure 3. Deprotonation mechanism according to the literature for myricetin, quercetin and kaempferol in aqueous medium (A). Resonance forms for the structures B and C (B).

Figure 4. 3D structures of the two PARAFAC components (first to the left and second to the right) obtained by multiplying the corresponding vectors (A). Correlation between concentrations obtained by the LC method and the score values obtained in PARAFAC model (B).

Figure 5. 2D representation of PARAFAC scores corresponding to the three components optimized model. Samples are represented according their origin. Score values for the first component against score values for the second component (A). Score values for the first component against score values for the third component (B).

Figure 6. Correlation between LC concentrations of quercetin and U-PLS/RBL (A) and N-PLS/RBL (B), respectively.

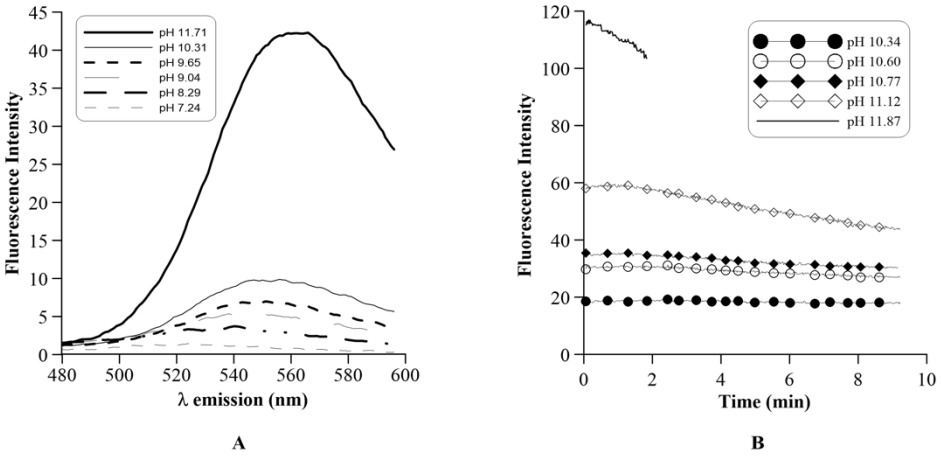


Figure 1

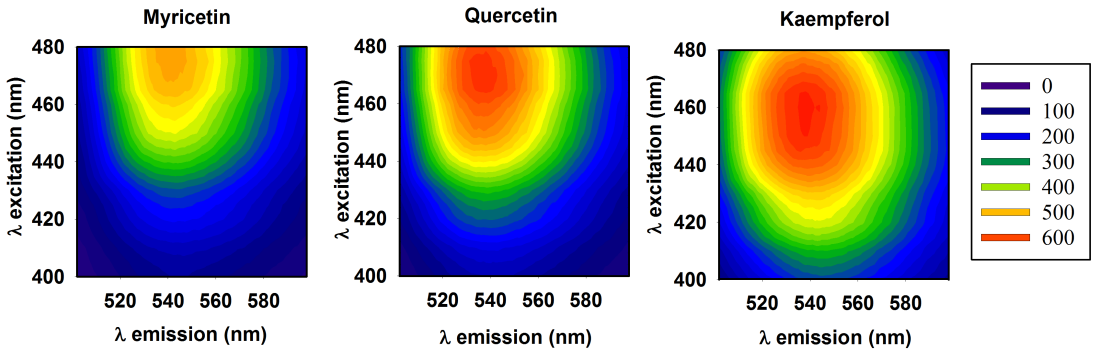
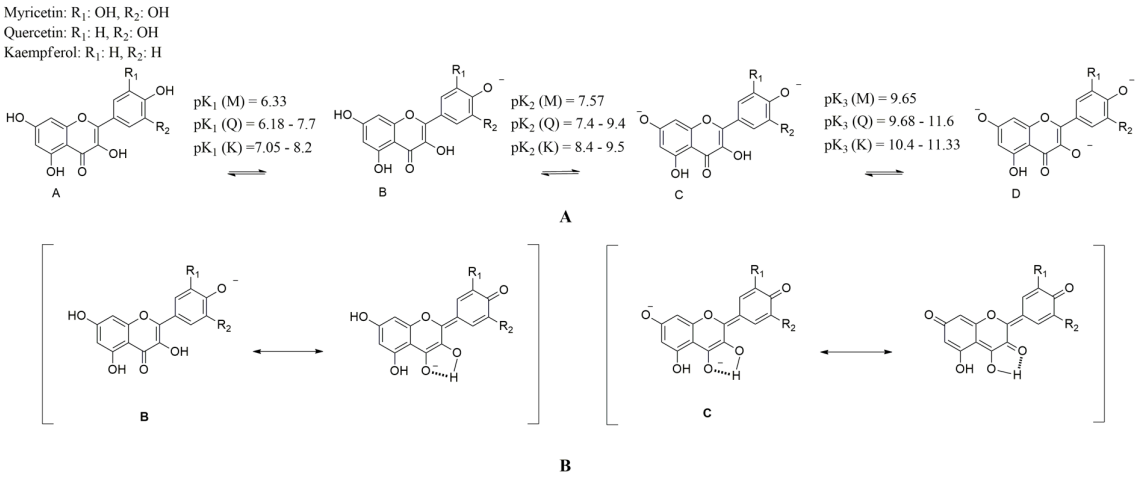
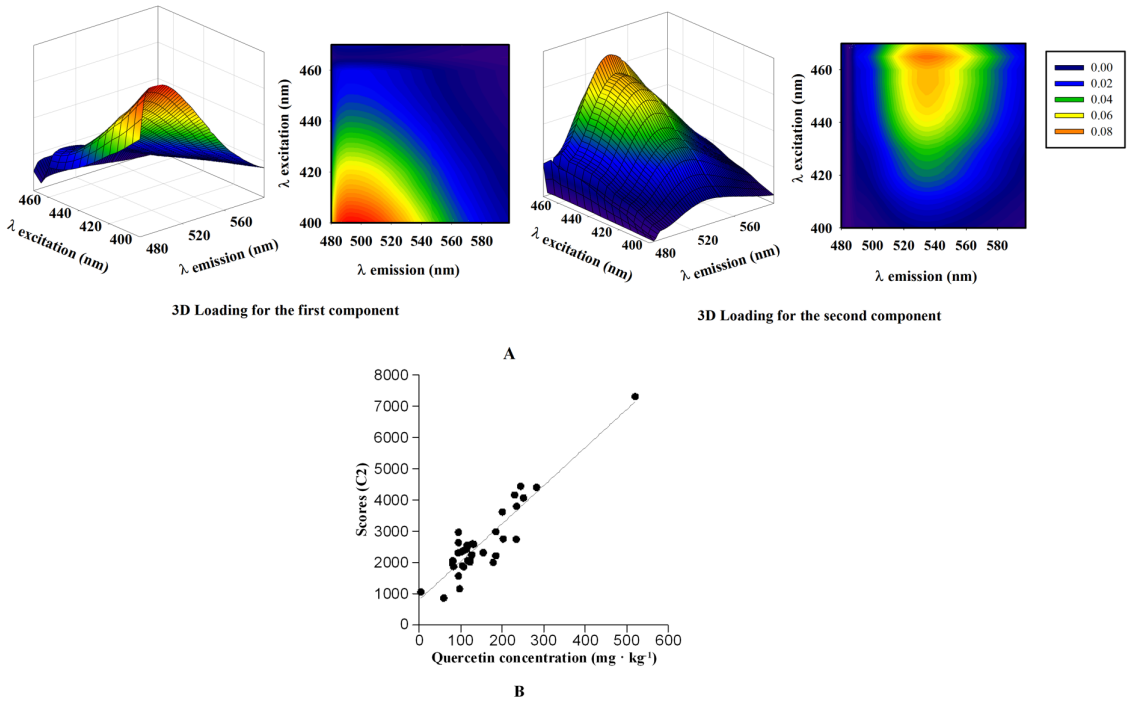


Figure 2



552

553 Figure 3



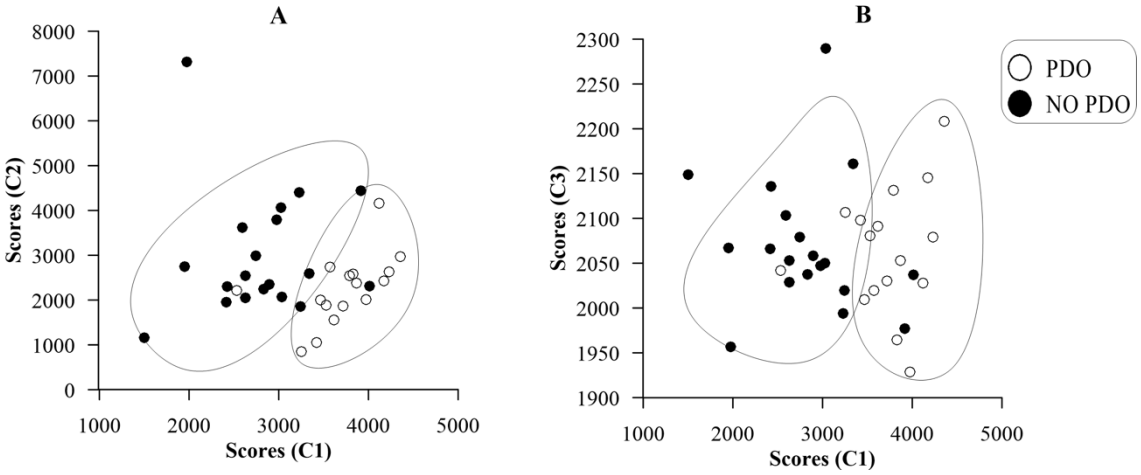
554

555 Figure 4

556

557

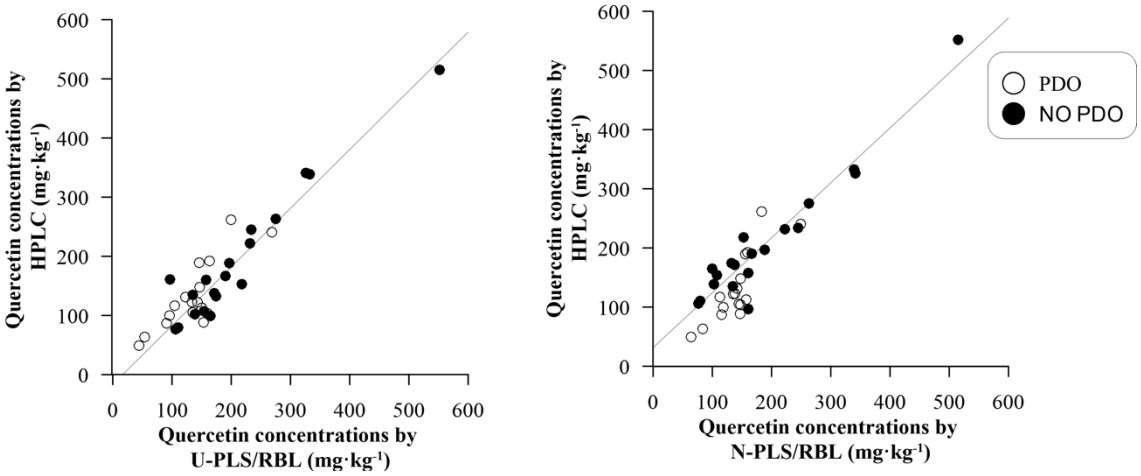
558



559

560

561 Figure 5



562

563

564 Figure 6