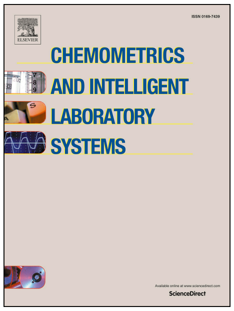

Impact of the pretreatment of ATR-FTIR signals on the figures of merit when PLS is used

C. Mas, L. Rubio, L. Valverde-Som, L.A. Sarabia, M.C. Ortiz

PII: S0169-7439(19)30816-0

DOI: https://doi.org/10.1016/j.chemolab.2020.104006

Reference: CHEMOM 104006

To appear in: *Chemometrics and Intelligent Laboratory Systems*

Received Date: 13 December 2019

Revised Date: 17 March 2020

Accepted Date: 19 March 2020

Please cite this article as: C. Mas, L. Rubio, L. Valverde-Som, L.A. Sarabia, M.C. Ortiz, Impact of the pretreatment of ATR-FTIR signals on the figures of merit when PLS is used, *Chemometrics and Intelligent Laboratory Systems* (2020), doi: https://doi.org/10.1016/j.chemolab.2020.104006.

Credit Author Statement:

C. Mas: Analytical methodology

L. Rubio: Investigation, Analytical Methodology Writing - Original Draft, Writing, Supervision

L. Valverde-Som: Investigation, Analytical Methodology, Supervision, Writing - Review & Editing

L.A. Sarabia Conceptualization, Methodology,Formal analysis, Software, Supervision, Writing - Review & Editing

M.C. Ortiz: Conceptualization, Supervision, Writing - Review & Editing

IMPACT OF THE PRETREATMENT OF ATR-FTIR SIGNALS ON THE FIGURES OF MERIT WHEN PLS IS USED

C. Mas[a], L. Rubio[a], L. Valverde-Som[a], L.A. Sarabia[b], M.C. Ortiz[a,1]

[a]*Department of Chemistry,* [b]*Department of Mathematics and Computation*

Faculty of Sciences, Universidad de Burgos

Plaza Misael Bañuelos s/n, 09001 Burgos (Spain)

Abbreviations[2]

Abstract

The efficiency of the analytical methods based on vibrational spectroscopy has been widely verified in a high number of publications. In addition, it has been recognized that the pretreatment of the original signals is absolutely necessary to obtain enough quality in the subsequent classification and/or regression models. In fact, an inappropriate pretreatment makes the results worse. It is also impossible to give "a priori" rules that guarantee the adequacy of a pretreatment for specific data.

The effect of the pretreatments is evaluated through their impact on the quality of the classification and /or regression models built from them due to the double dependence (on the data and on the purpose of the analysis). The effect of the pretreatment has been evaluated using partial least squares regression (PLSR) in some works and the root mean squares in prediction or in cross-validation has been always used as a criterion to evaluate the regression in all these cases. However, it seems appropriate to use quality criteria of the calibration of the analytical method through the figures of merit: the significance of the regression, the absence of constant or proportional bias, the residual standard deviation, the mean of the absolute values of the relative errors and the capability of detection.

In this work, the use of these analytical criteria in a desirability function is proposed for the first time with calibration data of oxybenzone obtained by ATR-FTIR and PLSR.

---

[1] Corresponding author. Telephone number: +34-947-259571. *E-mail address*: mcortiz@ubu.es (M.C. Ortiz).

[2] Attenuated total reflectance-Fourier transform infrared (ATR-FTIR), oxybenzone (BP3), capability of detection (CCβ) data preprocessing (DP), design of experiments (DoE), direct orthogonal signal correction (DOSC), extended multiplicative scatter correction (EMSC), infrared (IR), latent variable (LV), multiplicative scatter correction (MSC), near infrared (NIR), orthogonal signal correction (OSC), partial least squares regression (PLSR), root mean square error (RMSEC), root mean square error of cross-validation (RMSEC_CV), Standard error in prediction (SEP), standard normal variate (SNV).

This desirability function enables to choose the best pretreatment among the 39 possibilities studied. In addition, it is shown that the same optimum is not obtained if the minimum of RMSEC_CV is considered as a criterion.

Keywords

Pretreatment data; ATR-FTIR; PLS; oxybenzone; desirability function; figures of merit

1. Introduction

Data preprocessing (DP) has been recognized as a critical stage for high dimensional data analysis, particularly for data obtained using vibrational spectroscopy techniques. There are a lot of scientific works about different DP methods which can be classified in two large groups: i) pretreatment by means of orthogonal projections to extract the relevant information of a training set [1] using pure spectra and information extracted from experimental design and calibration datasets that includes the Orthogonal Signal Correction (OSC) family; ii) pretreatment by transformation of the spectral signal removing undesired physical phenomena [2,3]. This last group can be divided into two categories. The first one, scatter correction, includes Multiplicative Scatter Correction (MSC), Inverse MSC, Extended MSC (EMSC) [4,5], Extended Inverse MSC, normalization, and Standard Normal Variate (SNV). The second group belongs to the spectral derivatization group which includes the Norris-Williams derivatives and Savitzky-Golay polynomial derivatives. The relation between MSC and SNV [6] has been studied because are the pretreatments most employed in practice, also their effect on the signals [7] and the chemical interpretability of pretreated spectra [8]. At the same time, there are different review articles [5,9] and many tutorials that have been published to show the practical aspects of using DP, for example with near infrared (NIR) [2], Raman and infrared (IR) [3,5,10,11] or attenuated total reflectance-Fourier transform infrared (ATR-FTIR) [12] data.

However, despite the emphasis on the importance of the pretreatment of spectra data, there are few contributions on the search of the optimal pretreatment for a data set. The critical revision of the most common three strategies used for the data pretreatment concludes that *"all three have serious drawbacks: they may be time-consuming beyond practicability or may provide misleading results"* [13]. Then, it is necessary to consider jointly successive types of pretreatment. For this reason, J. Engel *et al.* [13] have used a search space formed by 7x10x10x7 pretreatments for baseline, scatter, noise, smoothing and scaling/transformations, respectively. Other 7 pretreatments are added for OSC and Direct Orthogonal Signal Correction (DOSC). A similar strategy includes binning, smoothing, normalization and baseline correction with 3528 different pretreatments that could be partially permuted [14]. In other work, experimental design (DoE), specifically a full factorial (2^4), is used as a strategy to explore the effect of these same types of pretreatments [15]. In reference [16] this strategy for pretreatment is linked with variable selection to improve the interpretation and the predictive capacity of a regression model. On the other hand, a factorial design 2^4 is used with three different methods applied in variable selection to the four pretreatments selected in [17]. A D-optimal design to optimize the pretreatment in three intermediate stages is also used for a metabolomic data fusion of urine and serum in [18].

Many works analyse or optimise the data pretreatment before applying a classification technique. In general, a criterion to evaluate the classifier (e.g. the % of misclassification) is used. Only H.J. Butler *et al.* [14] used several criteria combined additively in one unique value.

The need for carrying out a pretreatment of the spectra provided by vibrational spectroscopic techniques before applying a regression method, which is usually a Partial Least Squares Regression (PLSR), is completely accepted. In the literature consulted, the effect of changing the pretreatment is analysed through the change caused in the PLSR quality parameters. In this case, Root Mean Square Error of Prediction (RMSEP) [15,17], Root Mean Square Error of Cross-Validation (RMSEC_CV) [16] and a estimation of Root Mean Square Error (RMSEC) based on bootstrap [2] are used. In these works, the use of quality criteria of an analytical method evaluated by means of the regression “predicted concentration with PLSR *versus* true concentration of the test sample”, which is named as the accuracy line in prediction, is not tackled. This work proposes to study the effect of the pretreatment on the quality parameters of the PLSR model and on the accuracy line in prediction by means of ATR-FTIR data of the determination of oxybenzone, benzophenone-3 (BP3). BP3 is an additive used in the manufacture of sunscreen cosmetic creams and its determination by ATR-FTIR is possible due to the amount found in creams.

2. Material and methods

2.1 Chemicals

2-hydroxy-4-methoxybenzophenone (oxybenzone, CAS no. 131-57-7, 98% purity) was purchased from Sigma-Aldrich (Steinheim, Germany).

Ethanol (96% vol., CAS no. 64-17-5, HiPerSolv CHROMANORM®, gradient grade for HPLC) was supplied by VWR International (Radnor, Pennsylvania, USA) and acetone (CAS no. 67-64-1) for liquid chromatography Lichrosolv® was from Merck KGaA (Darmstadt, Germany).

2.2 Standard solutions

A stock solution of BP3 at 30275 mg L^{-1} was prepared in ethanol and intermediate solutions at concentrations of 5000, 10000, 15000 and 20000 mg L^{-1} were prepared from that stock solution in ethanol as calibration standards. All solutions, whose weights were controlled to verify that the solvent had not evaporated, were stored in crimp vials at 4ºC and protected from light. The stability of the solutions of BP3 has been verified by GC/MS, being the stock and intermediate solutions of BP3 stable for 15 days as can be seen in ref. [19].

2.3. Instrumental

An Agilent Cary 630 FTIR spectrometer coupled to a ZnSe ATR module for measuring liquids samples (Agilent Technologies, Santa Clara, CA, USA) was used to perform the analyses. The number of reflections of the crystal with this module was 5. The main optical unit dimensions were only 16 × 22 × 13 cm. The Cary 630 FTIR spectrometer contained a unique Michelson interferometer.

The spectral range to collect the absorbance signal was from 650 to 4000 cm^{-1} with 32 scans and the spectral resolution was fixed to 4 cm^{-1} using Happ-Genzel as apodization function [20]. The method gain was set at 255.

2.4. Steps to measure the sample

First, air collection as background was selected with 16 scans. The background was collected at the beginning of the sequence and it was not measured again between samples. Then, a volume of 100 µL of the sample was put on the crystal. In this step, the absorbance signal was collected. Finally, two solvents (ethanol and acetone) were used to clean the crystal using delicate task wipers (Kimwipes®, from Sigma-Aldrich). A volume of 100 µL of ethanol was put on the crystal and the signal was recorded to check if the crystal was clean.

2.5. Software

MicroLab PC, version 5.3.1748 (Agilent Technologies, Inc.) with Data Analysis software was used for acquiring data. The different pretreatments of the signals and PLSR models were performed with the PLS_Toolbox [21] used under MATLAB environment [22]. The regression models were fitted and validated using STATGRAPHICS Centurion XVI [23]. Capability of detection (CCβ) was calculated using the DETARCHI program [24].

3. Calculations

3.1. Data

A wide calibration range and few concentration levels were used as in routine analyses. The calibration set was made up of four samples at concentrations of 5000, 10000, 15000 and 20000 mg L^{-1} of BP3 and a blank. The test set was made up of other samples of the same concentrations of BP3 which were prepared again and independently of the calibration standards. The matrix dimensions in both cases are 5×1798.

3.2. Pretreatment

The introduction section described different data pretreatments. In this work, the most common data pretreatment is chosen for ATR-FTIR spectroscopy data with liquid samples:

1) Multiplicative Scatter Correction (MSC) or Standard Normal Variate (SNV) to correct the scattering signal.

2) Savitzky-Golay smoothing (SG) with windows of 15 (or 25) points, a second order polynomial and second derivative.

3) Mean centering or Autoscaling for scale correction.

In addition, the possibility of applying only one, two or no correction has been included. Scaling is the only pretreatment of the three ones described above that is not directly related to the elimination of any signal artefact. If scaling is used, it should be the last step of the pretreatment because its effect is to unify the statistical scale of the variables (mean and/or variances). Therefore, if scaling is used prior to scatter correction, the scaling effect will be partially attenuated by the scatter correction method. However, the other two steps may be changed in order. This may be important, e.g., the effect of exchanging the MSC and SG pretreatments is checked in ref. [25]. Table 1 shows the code of the different pretreatments used. By way of example, the code "253" means that SNV has been applied in the first step, Savitzky-Golay smoothing with windows of 15 points in the second step, and Autoscaling in the last step. The code “141” means that no pretreatment has been performed. When the change of order between scatter and smoothing is considered, duplication must be avoided when one of them is not applied. By way of example, the pretreatment “153” (which means using the code of Table 1: do not apply scatter correction, SG with 15 points and autoscaling) would be the same pretreatment as "513". Therefore, only a total of 39 different pretreatments have been applied.

Given a set of spectra S_i, i=1,.., n (e.g. a set of calibration samples), ref. [6] shows that there is a linear relationship between each spectrum transformed by SNV and the one transformed by MSC, although that linear relationship is different for each S_i. Therefore, the pretreated spectra are similar except for a rotation and an offset correction and it is generally assumed that “MSC and SNV are the same for the most practical applications” [9,26] but the results are not always very similar. Fearn et al. [7] obtained “very different results when NIR spectral data were pre-treated with SNV and MSC, the former leading to a striking ellipsoidal structure in a plot of principal component scores and the latter to a plot with many extreme outliers”. In addition, ref. [8] shows that MSC (and other pretreatments of the same family including non-linear approaches as EMSC) produces a shift along the signal profile which leads to artefacts in the principal component analysis of the spectra. Both pretreatments have been included since the aim of this work is the analysis of the effect of the pretreatments on the PLSR calibrations and there are no comparative studies in literature about the differences caused by SNV and MSC in the figures of merit.

<Table 1>

3.3. PLSR model

The PLSR model is built with the training data set and each of the 39 different pretreatments. The number of latent variables is the one that provides the minimum RMSEC_CV with leave-one-out as cross-validation. Then, the model is applied to the test set. In both cases, Hotelling's T^2 and Q-residual statistic at a 95% confidence level are used to check if there is any outlier. No outlier data was found in the built models.

4. Results and discussion

4.1. Effect of the pretreatment on PLSR models

Table 2 shows all the pretreatments with different information about the PLSR models. The second column of this table indicates the number of latent variables in the model, whereas root mean square in calibration (RMSEC) and in cross-validation (RMSEC_CV) are collected in the third and fourth columns, respectively. In addition, standard error in prediction (SEP) has been calculated.

<Table 2>

The pretreatments with Savitzky-Golay smoothing with windows of 15 points and without scaling (codes 151, 251 and 351) show the greatest values of RMSEC_CV, SEP and RMSEC. On the other hand, the pretreatments without Savitzky-Golay smoothing and no autoscaling (codes 342, 341, 241, 242, 142 and 141) show the lowest values of RMSEC_CV with values between 2881.7 and 3002.9, but intermediate values of RMSEC (between 900 and 1000). In those cases, the selection of scatter correction (1, 2 or 3 in the first position of the code) is indifferent. The effect of no scaling or mean centering (1 or 2 in the third position of the code) is very small and it is only observed when the scatter correction is used, in which case mean centering provides better values of RMSEC, RMSEC_CV and SEP, but increases ERROR, s_{yx}, and CCβ. The absolute minimum is achieved with a value of 2881.7 of RMSEC_CV with the pretreatment 342.

4.2. Effect of the pretreatment on the figures of merit of the analytical method

In the scope of chemical analysis, it is important to validate the calibration model of an analytical method as well as to obtain the figures of merit. The statistical procedure for this task can be found in reference [27]. It is interesting to know, through the PLSR model, the impact on these figures of merit to evaluate the different pretreatments carried out. First, the PLSR model is built; then, the regression line "predicted concentration with PLSR *versus* true concentration of BP3" which is known as accuracy line is validated. The test set samples of these regressions were not used to build the previous PLSR model.

Table 2 shows the parameters of the accuracy lines in columns 6, 7, 8 and 9 such as the correlation coefficient (r), residual standard deviation (s_{yx}), intercept (b_0) and slope (b_1), respectively.

4.2.1. Significance of the accuracy line

The hypothesis test to evaluate the significance of the accuracy line involves:

> H_0: the regression does not explain a variance bigger than the residual.
>
> H_a: the regression explains the variability of the response, so the regression is significant.

The null hypothesis (H_0) is rejected if the p-value of test is lower than 0.05, in which case the alternative hypothesis (H_a) is accepted. Table 2 shows the p-values of the significance test (p-s, column 10) which are significant and H_0 is rejected. These results are coherent with the correlation coefficients which have values between 0.96 and 1.00 (Table 2, column 6).

4.2.2. Bias of the accuracy line

The accuracy line should have the intercept equal to zero and the slope equal to one. Otherwise, the analytical procedure has constant and/or proportional bias.

The hypothesis test is the following one (page 141 of ref. [27]):

> H_0: intercept and slope are equal to zero and one, respectively (b_0=0 and b_1=1).
>
> H_a: is not the case, the accuracy line is biased.

The null hypothesis (H_0) is rejected if the p-value is lower than 0.05, in which case the alternative hypothesis (H_a) is accepted. Table 2 shows the p-values (p-F, column 11). The nine pretreatments with Savitzky-Golay smoothing with windows of 15 points (all of them marked with number 5 in the second position of the code in PRET of Table 2) show values lower than 0.05, therefore, H_0 is rejected.

These pretreatments (151, 152, 153, 251, 252, 253, 351, 352 and 353) are not considered acceptable from an analytical point of view; in addition, the greatest values of RMSEC_CV and SEP are among them.

4.2.3. Relative error

All the pretreatments (39) are evaluated by the mean of the absolute value of the relative error, ERROR (see Table 2, column 12) which is calculated through the concentration of BP3. The lowest value of ERROR (11.1%) is reached with the pretreatments 631, 532 and 531. In the three cases, a noise correction (Savitzky-Golay smoothing with windows of 15 or 25 points) is applied first, then a scatter correction (MSC) is considered and, finally, a mean centering in one of them. These three pretreatments have nearly the same value, 4742, of RMSEC_CV in the upper half of the range of values of this parameter. In addition, the value of SEP is 2881 in the three cases and 18 pretreatments have a value of SEP lower than this one.

4.2.4. Capability of detection

Finally, the capability of detection (CCβ) is another analytical criterion which is defined by ISO 11843 Part 1 and 2 [28,29] and IUPAC [30] as the smallest concentration of the substance that may be detected, x_d, with the probabilities of false positive (α) and false negative (β) fixed at 0.05. *"Capability of detection is the smallest value of the net state variables which can be detected with a probability of 1 – β as different from zero"*. This parameter is considered as the hypothesis test, where α is the probability of false positive and β is the probability of false negative. ISO 11843 defines x_d which is equivalent to CCβ in IUPAC. In this specific case, CCβ is the amount of BP3 such that there is a probability of α and β of 0.05 to get a false positive and false negative result, respectively. Table 2 (column 13) contains the CCβ values and shows that the pretreatment has an important effect, since the values ranged from 88 to 23062 mg L^{-1}. Therefore, the pretreatments 533, 633, 263 and 622 provide the minimum values.

4.3. Visualizing the conflict caused by the pretreatments

4.3.1. Parallel coordinates

The representation of the parallel coordinates of all columns in Table 2 enables a joint interpretation of the results and the visualization of the conflict with all the criteria caused by different pretreatments of the ATR-FTIR signals.

The vertical axes, Figure 1, show the variables of column 3 to 13 of Table 2 which have been described in sections 4.1 and 4.2. The ranges of the variables are very different, so they have been scaled between 0 to 1; otherwise it would be impossible to visualize the individual values of each of them. However, the corresponding minimum and maximum values are indicated at the end of each axis to show the relative position of each represented value. The values obtained in each pretreatment for each of the 11 variables are joined by a polygonal line. The grey lines show the 9 pretreatments that led to biased calibrations in section 4.2.2.

The continuous red line shows the pretreatments (341 and 342) that make RMSEC_CV minimum. Both lines coincide in all the variables. The dashed red lines correspond to the pretreatments (143 and 162) that make SEP minimum which differ in RMSEC_CV and ERROR and less in s_{yx} and CCβ.

The overlapping green lines show three pretreatments not biased and with less ERROR (section 4.2.3). These pretreatments (531, 532 and 631) have RMSEC_CV values far from the minimum. The same goes for the corresponding values of s_{yx} and CCβ. Therefore, the pretreatments that minimize the mean of the absolute values of the relative errors in prediction do not provide minimum values of RMSEC_CV or SEP.

The magenta lines highlight four pretreatments with the lowest CCβ (better capability of detection) values (section 4.2.4). These pretreatments (263, 622, 533 and 633) do not have minimum values of RMSEC_CV or SEP. However, these pretreatments reach the minimum value of s_{yx} but not of the ERROR since they take different values in that variable.

From the analysis of the data of Table 2 and of its graphical version in Figure 1, it can be concluded that: (1) minimizing the quality criteria of the PLS regression (RMSEC_CV and SEP) does not lead to minimize the analytical quality criteria (s_{yx}, ERROR and CCβ), and (2) a minimum of the three analytical quality criteria is not obtained simultaneously.

Figure 1 shows other general aspects, for example, the general contradictory behaviour between RMSEC and RMSEC_CV. The black line in Figure 1 corresponds to no pretreatment of the data (code 141) and it is observed that most of the pretreatments make RMSEC_CV worse as already explained [13,15]. The same effect can be observed in SEP even in s_{yx}, ERROR and CCβ but in a lower degree in these last cases.

<Figure 1>

4.3.2 Principal component analysis

It has been proved that the pretreatments that provide the minimum value of each variable do not share any pattern or characteristic. To describe globally the behaviour of the variables with regard to the pretreatments, a principal component analysis of RMSEC, RMSEC_CV, SEP, r, s_{yx}, b_0, b_1, ERROR and CCβ is useful. Once the variables have been autoscaled, a 87.9 % of the variance is explained with 3 PC. The evolution of the eigenvalues and the percentage of explained variance are contained in Table 3.

<Table 3>

The loadings of the original variables in the first three components are displayed in Figure 2. As can be seen in Figure 2a, the variables related to the accuracy line (ERROR, s_{yx} and CCβ) have little correlation and the same happens with RMSEC with respect to RMSEC_CV and SEP. Therefore, there is no direction in the space of the principal components linked exclusively to one type of the variables, either those that define the quality of PLSR, or those that define the analytical quality of the method.

Figure 2b contains the loadings of the first and third component. As can be seen in this last figure, RMSEC is not related to the rest of the variables in this projection plane. To sum up, when the variability of the different pretreatments is studied, it is necessary to consider the joint contribution of the variables s_{yx}, ERROR and CCβ.

<Figure 2>

4.4 A Multicriteria solution

The quality criteria of the analytical method analysed in section 4.2 are two p-values corresponding to the significance tests of the calibration line (p-s), and the joint one for the intercept and of the slope (p-F). On the other hand, the rest of the criteria are the correlation coefficient (r), the residual standard deviation (s_{yx}), the intercept (b_0), the slope (b_1), the mean of the absolute value of the relative errors (ERROR) and the capability of detection (CCβ).

Section 4.3 has shown that it is not possible to find a pretreatment that optimize all of them simultaneously. Therefore, a desirability function of Derringer [31] built with these criteria is proposed to minimize s_{yx}, ERROR and CCβ and it is also combined with the conclusion of the two hypothesis tests.

The hypothesis tests are a very useful tool in chemical analysis [32] but the p-values cannot be used quantitatively as an index of the degree of validity of the null or alternative hypothesis [33]. Therefore, only the result of the test will be considered instead of taking into account the p-values (p-s and p-F), that is, if the null hypothesis is rejected or accepted. Therefore, two binary variables, d_i i=1,2, will be assigned. A significance level of α =0.05 is considered as usual for the two tests as in section 4.2.1 and 4.2.2. In particular, d_1 will have the value of 1 if p-s is lower than α. d_2 will have the value of 1 if p-F is greater than α. If the accuracy line is not significant or there is bias, the product of these functions will be zero.

The individual desirability functions d_3, d_4 and d_5 to minimize s_{yx}, ERROR and CCβ, respectively, are defined in the same way. d_i=1 if the value of the variable is below the 10th percentile of the data, d_i=0 if the value of the variable is over the 75th percentile of the data and d_i is a linear function between both values if the value of the variable is between them. The choice of the thresholds (75th and 10th percentiles) is subjective and the value of the desirability is modified when they are changed but not the order, in particular, the pretreatment in which the minimum and maximum are reached.

Finally, each pretreatment is evaluated with the global desirability function $D = d_1 \times d_2 \times d_3^{1/3} \times d_4^{1/3} \times d_5^{1/3}$. The weighting of the three figures of merit is subjective, but it does not seem conceptually that one of them is more important than the other ones in this problem. ERROR and CCβ have a direct analytical meaning, but s_{yx} affects the accuracy in the determination. To sum up, a pretreatment will have desirability equal to zero if it causes a biased or non-significant accuracy line and if one or more of the s_{yx}, ERROR and CCβ values are in the corresponding upper quartile.

The result is shown in Table 4, being 622 (S-G smoothing, window 25 pt/SNV scatter correction/Mean centering) the pretreatment which optimizes the figures of merit of the method. On the other hand, the pretreatments that provide the lowest values of RMSEC_CV (342, 341, 241, 242, 142 and 141, section 4.1) have a value of global desirability equal to zero since they are in the upper quartile in at least one of the figures of merit (s_{yx}, ERROR and CCβ). The 9 pretreatments that lead to a biased accuracy line have also desirability equal to zero (section 4.2.2). The five pretreatments with a value of desirability greater than or equal to 0.90 (622, 632, 163, 263 and 363) share a smoothing with windows of 25 points which is more effective if it is done first. Regarding the scatter correction, it can be seen that the SNV pretreatment gives lower global desirability values than MSC when it is applied after smoothing and keeping the rest of the pretreatments the same. This tendency is reversed when the correction of the dispersion is applied first.

<table 4>

5. Conclusions

The effect of 39 different pretreatments which combined scatter, noise and scaling correction on ATR-FTIR signals for the determination of oxybenzone was studied.

The pretreatments behave in a different way if the evaluation is about the quality of the PLSR models or over the quality parameters of the analytical procedure. In addition, some PLSR models provide biased results in the accuracy line, so it is important to consider both evaluations to select the best pretreatment.

The proposal is to work jointly with the quantitative criteria of the calibration in prediction (mean of absolute values of relative errors, residual standard deviation, and capability of detection) by means of a desirability function which also includes the significance of the regression and the absence of bias.

The procedure could be applied to other calibration independently of the multivariate regression model because the criteria depend on the accuracy line.

6. Acknowledgements

The authors thank the financial support provided by Spanish MINECO (AEI/FEDER, UE) through project CTQ2017□88894□R and by Junta de Castilla y León through project BU012P17 (both co□financed with European FEDER funds). L. Valverde-Som thanks JCyL for her postdoctoral contract through BU012P17 project.

7. References

[1] J.C. Boulet, J.M. Roger, Pretreatments by means of orthogonal projections, Chemometr. Intell. Lab. 117 (2012) 61–69. http://doi.org/10.1016/j.chemolab.2012.02.002.

[2] A. Rinnan, F. van den Berg, S.B. Engelsen, Review of the most common pre-processing techniques for near-infrared spectra, TRAC-Trend. Anal. Chem. 28 (2009) 1201-1222. https://doi.org/10.1016/j.trac.2009.07.007.

[3] R. Gautam, S. Vanga, F. Ariese, S. Umapathy, Review of multidimensional data processing approaches for Raman and infrared spectroscopy, EPJ Techniques and Instrumentation (2015) 2:8. https://doi.org/10.1140/epjti/s40485-015-0018-6.

[4] N.K. Afseth, A. Kohler, Extended multiplicative signal correction in vibrational spectroscopy, a tutorial, Chemometr. Intell. Lab. 117 (2012) 92–99. https://doi.org/10.1016/j.chemolab.2012.03.004.

[5] M. Féré, O. Piot, L.H Liu, A. Beljebbar, V. Untereiner, D. Gheldof, M. Chollat, J. experimental and instrumental conditions, Vib. Spectrosc. 103 (2019) 102931. https://doi.org/10.1016/j.vibspec.2019.102931.

[6] M.S. Dhanoa, S.J. Lister, R. Sanderson, R.J. Barnes, The link between Multiplicative Scatter Correction (MSC) and Standard Normal Variate (SNV) transformations of NIR spectra, J. Near Infrared Spec. 2 (1994) 43-47. https://doi.org/10.1255/jnirs.30.

[7] T. Fearn, C. Riccioli, A. Garrido-Varo, J.E. Guerrero-Ginel, On the geometry of SNV and MSC, Chemometr. Intell. Lab. 96 (2009) 22–26. https://doi.org/10.1016/j.chemolab.2008.11.006.

[8] P. Oliveri, C. Malegori, R. Simonetti, M. Casale, The impact of signal pre-processing on the final interpretation of analytical outcomes-A tutorial, Anal. Chim. Acta 1058, (2019), 9-17. https://doi.org/10.1016/j.aca.2018.10.055.

[9] A. Rinnan, Pre-processing in vibrational spectroscopy – when, why and how, Anal. Methods-UK 6 (2014) 7124-7129. https://doi.org/10.1039/c3ay42270d.

[10] P. Heraud, B.R. Wood, J. Beardall, D. McNaughton, Effects of pre-processing of Raman spectra on *in vivo* classification of nutrient status of microalga cells, J. Chemometr. 20 (2006) 193–197. https://doi.org/10.1002/cem.990.

[11] H. J. Byrne, P. Knief, M.E. Keating, F. Bonnier, Spectral pre and post processing for infrared and Raman spectroscopy of biological tissues and cells, Chem. Soc. Rev. 45 (2016) 1865-1878. https://doi.org/10.1039/c5cs00440c.

[12] L.C. Lee, C.Y. Liong, A. A. Jemain, A contemporary review on Data Preprocessing (DP) practice strategy in ATR-FTIR spectrum, Chemometr. Intell. Lab. 163 (2017) 64–75. https://doi.org/10.1016/j.chemolab.2017.02.008.

[13] J. Engel, J. Gerretzen, E. Szymanska, J.J. Jansen, G. Downey, L. Blanchet, L.M.C. Buydens, Breaking with trends in pre-processing?, TRAC-Trend. Anal. Chem. 50 (2013) 96–106. https://doi.org/10.1016/j.trac.2013.04.015.

[14] H.J. Butler, B.R. Smith, R. Fritzsch, P. Radhakrishnan, D.S. Palmer, M.J. Baker, Optimised spectral pre-processing for discrimination of biofluids *via* ATR-FTIR spectroscopy, Analyst 143 (2018) 6121-6134. https://doi.org/10.1039/c8an01384e.

[15] J. Gerretzen, E. Szymańska, J.J. Jansen, J. Bart, H.J. van Manen, E.R. van den Heuvel, L.M.C. Buydens, Simple and effective way for data preprocessing selection based on design of experiments, Anal. Chem. 87 (2015) 12096–12103. https://doi.org/10.1021/acs.analchem.5b02832.

[16] J. Gerretzen, E. Szymanska, J. Bart, A.N. Davies, H.J. van Manen, E.R. van den Heuvel, J.J. Jansen, L.M.C. Buydens, Boosting model performance and interpretation by entangling preprocessing selection and variable selection, Anal. Chim. Act 938 (2016) 44-52. https://doi.org/10.1016/j.aca.2016.08.022.

[17] L. Li, S. Quan, D. Li, J. Wang, H. Zang, L. Zhang, Development of near infrared spectroscopy methodology for human albumin determination using a new calibration approach, Spectrochim. Acta A 217 (2019) 256–262. https://doi.org/10.1016/j.saa.2019.03.100.

[18] H. Zheng, A. Cai, Q. Zhou, P. Xu, L. Zhao, C. Li, B. Dong, H. Gao, Optimal preprocessing of serum and urine metabolomic data fusion for staging prostate cancer through design of experiment, Anal. Chim. Act 991 (2017) 68-75. https://doi.org/10.1016/j.aca.2017.09.019.

[19] L. Rubio, L. Valverde-Som, L.A. Sarabia, M.C. Ortiz, Improvement in the identification and quantification of UV filters and additives in sunscreen cosmetic creams by gas chromatography/mass spectrometry through three-way calibration techniques, Talanta (2019) 120156. https://doi.org/10.1016/j.talanta.2019.120156.

[20] R.S. Bretzlaff, T.B. Bahder, Apodization effects in Fourier transform infrared difference spectra, Revue Phys. Appl. 21 (1986) 833-844. https://doi.org/10.1051/rphysap:019860021012083300.

[21] B.M. Wise, N.B. Gallagher, R. Bro, J.M. Shaver, W. Winding, R.S. Koch, PLS Toolbox 8.5.2, Eigenvector Research Inc., Wenatchee, WA, USA.

[22] MATLAB, version 9.3.0.713579 (R2017b), The Mathworks, Inc., Natick, MA, USA.

[23] STATGRAPHICS Centurion XVI Version 16.1.05 (32 bit), Statpoint Technologies, Inc., Herndon, VA, USA, 2010.

[24] L.A. Sarabia, M.C. Ortiz, DETARCHI. A program for detection limits with specified assurance probabilities and characteristic curves of detection, Trends Anal.Chem.13 (1994) 1-6. https://doi.org/10.1016/0165-9936(94)85052-6.

[25] H. Chen, Q. Song, G. Tang, Q. Feng, L. Lin, The Combined Optimization of Savitzky-Golay Smoothing and Multiplicative Scatter Correction for FT-NIR PLS Models, ISRN Spectroscopy (2013) Article ID 642190, 9 pages http://dx.doi.org/10.1155/2013/642190.

[26] T. Fearn, The effect of spectral pre-treatments on interpretation, NIR News 20 (2009) 15.16. 15e16, https://doi.org/10.1255/nirn.1146.

[27] M.C. Ortiz, M.S. Sánchez, L.A. Sarabia, Quality of analytical measurements: univariate regression, in: S.D. Brown, R. Tauler, B. Walczak (Eds.), Comprehensive chemometrics. Chemical and biochemical data analysis. Elsevier, v1, Amsterdam, 2009, pp. 127-169. https://doi.org/10.1016/B978-044452701-1.00091-0.

[28] International Organization for Standardization, ISO 11843-1:1997, Capability of detection - Part1: Terms and definitions, Genève, Switzerland, 1997.

[29] International Organization for Standardization, ISO 11843-2:2000, Capability of detection – Part2: Methodology in the linear calibration case, Genève, Switzerland, 2000.

[30] J. Inczédy, T. Lengyel, A.M. Ure, A. Gelencsér, A. Hulanicki, International Union of Pure and Applied Chemistry, IUPAC Compendium of analytical nomenclature, third ed., Baltimore, Port City Press Inc., 2000.

[31] G. Derringer, R. Suich, Simultaneous optimization of several response variables, J. Qual. Technol. 12 (1980) 214-219.

[32] M.C. Ortiz, L.A. Sarabia, M.S. Sánchez, Tutorial on evaluation of type I and type II errors in chemical analyses: From the analytical detection to authentication of products and process control, Anal. Chim. Acta 674 (2010) 123-142, https://doi.org/10.1016/j.aca.2010.06.026

[33] R.G. Brereton, The use and misuse of p values and related concepts, Chemom. Intell. Lab. Syst. 195 (2019) 103884 https://doi.org/10.1016/j.chemolab.2019.103884

FIGURE CAPTIONS

Figure 1 Parallel coordinates plot of the values of the quality variables of the PLSR model and the accuracy line. The codification of variables is the same as in Table 2.

Figure 2 Principal component analysis of the variables RMSEC, RMSEC_CV, SEP, r, s_{yx}, b_0, b_1, ERROR and CCβ of Table 2. A) Loadings on the first and second PC and B) loadings on the first and third PC.

Table 1 Pretreatments and their codification.

Code / Scatter correction		Code / Noise correction		Code / Scaling	
1	No scatter correction	4	No noise removal	1	No scaling
2	SNV	5	S-G smoothing, window 15 points	2	Mean centering
3	MSC	6	S-G smoothing, window 25 points	3	Autoscaling

SNV, Standard Normal Variate; MSC, Multiplicative Scatter Correction; S-G, Savitzky-Golay.

Table 2 PLS models and accuracy line (see the pretreatments codification in Table 1)

PRET	L.V.	RMSEC	RMSEC_CV	SEP	r	s_{yx}	b_0	b_1	p-s	p-F	ERROR	CCβ
141	3	994.02	3002.91	2422.31	0.9709	1703.7	4020	0.757	0.00590	0.118	24.4	12687
142	2	990.57	2967.13	2367.66	0.9705	1746.1	3976	0.770	0.00610	0.126	24.9	12790
143	2	1098.44	3487.76	1776.56	0.9853	1310.0	2755	0.827	0.00210	0.154	16.0	8928
151	2	1779.22	6266.01	4965.47	0.9650	943.3	7150	0.380	0.00780	0.004	38.1	13975
152	2	167.99	5684.47	3880.48	0.9747	1096.7	5884	0.524	0.00480	0.013	33.8	11796
153	2	89.47	4343.81	3089.47	0.9803	1160.6	4476	0.630	0.00330	0.031	22.0	10376
161	2	912.70	4717.50	2375.71	0.9806	1346.0	3334	0.738	0.00320	0.100	11.8	10279
162	1	660.33	4061.57	1734.71	0.9885	1146.9	2509	0.820	0.00150	0.141	12.8	7883
163	2	738.88	3751.92	2276.04	1.0000	63.7	0	1.000	0.00001	1.000	14.5	359
241	3	899.14	2889.14	2272.84	0.9766	1542.7	3795	0.767	0.00430	0.108	22.7	11331
242	2	900.19	2889.48	2271.66	0.9766	1541.8	3794	0.768	0.00430	0.108	22.7	11323
243	2	865.46	3864.61	2480.90	0.9686	1762.8	4225	0.753	0.00660	0.113	27.4	13202
251	2	1740.41	6257.89	4939.87	0.9666	927.7	7101	0.383	0.00730	0.004	37.8	13637
252	2	105.20	5974.19	4230.66	0.9734	1024.5	6322	0.477	0.00520	0.008	34.2	12113
253	2	84.07	4316.10	3148.11	0.9797	1163.6	4534	0.623	0.00350	0.030	22.1	10533
261	2	928.66	4726.40	2361.94	0.9807	1348.1	3310	0.740	0.00320	0.103	11.5	10265
262	1	811.85	4480.37	2033.53	0.9863	1189.3	2851	0.778	0.00190	0.109	11.3	8614
263	2	20.53	3759.02	2462.69	1.000	26.50	0	1.000	0.00001	1.000	15.3	88
341	3	899.52	2881.97	2273.51	0.9767	1539.4	3801	0.767	0.00430	0.107	22.7	11312
342	2	899.62	2881.71	2273.52	0.9767	1539.4	3801	0.767	0.00430	0.107	22.7	11312
343	2	790.41	3861.69	2408.57	0.9714	1692.4	4140	0.758	0.00580	0.107	27.3	12581
351	2	1743.09	6254.58	4943.96	0.9667	925.6	7112	0.383	0.00730	0.004	37.8	13626
352	2	113.70	5969.39	4236.79	0.9726	1038.6	6340	0.476	0.00540	0.008	34.4	12292
353	2	84.28	4308.06	3151.74	0.9798	1157.6	4554	0.622	0.00340	0.029	22.2	10490
361	2	919.48	4720.54	2363.63	0.9808	1343.9	3319	0.740	0.00320	0.101	11.5	10240
362	2	53.54	4535.90	2110.54	0.9889	3103.8	3102	0.759	0.00140	0.501	13.7	23062
363	2	721.44	3755.32	2461.20	1.0000	63.7	0	1.000	0.00001	1.000	15.4	359
521	3	36.13	6145.40	4518.96	0.9767	46.64	0	1.000	0.00001	1.000	34.2	263
522	2	36.29	6144.19	4519.84	0.9767	46.85	0	1.000	0.00001	1.000	34.3	264
523	2	84.34	4383.23	3070.23	0.9999	108.89	1	1.000	0.00001	1.000	22.5	614
621	2	1679.93	6227.64	4906.01	0.9714	2105.97	522	0.944	0.00580	0.917	37.1	12581
622	2	74.31	4859.11	2353.86	1.0000	15.63	0	1.000	0.00001	1.000	12.5	88

623	2	34.57	6144.32	3104.53	0.9999	108.74	1	1.000	0.00001	1.000	22.3	613
531	2	958.31	4742.01	2355.93	0.9908	1225.76	185	0.982	0.00110	0.974	11.1	7039
532	2	958.60	4742.02	2355.93	0.9908	1226.13	184	0.982	0.00110	0.974	11.1	7041
533	2	12.10	3788.33	2577.24	1.0000	15.62	0	1.000	0.00001	1.000	17.6	88
631	2	957.54	4741.76	2355.06	0.9908	1224.79	181	0.982	0.00110	0.974	11.1	7033
632	2	74.41	4850.72	2345.98	0.9999	96.06	1	1.000	0.00001	1.000	12.6	542
633	2	12.12	3779.46	2530.69	1.0000	15.64	0	1.000	0.00001	1.000	17.5	88

PRET, spectra data pretreatment (see the codification in Table 1); LV, number of latent variable in the PLS model; RMSEC, Root Mean Squares Error in Calibration and the same in cross validation, RMSEC_CV; SEP, Standard Error in Prediction; r, correlation coefficient; s_{yx} residual standard deviation; b_0, intercept ; b_1, slope; p-s, p-value of the significance test; p-b_0&b_1, p-value to jointly test the intercept = 0 and slope = 1; ERROR (%), mean of the absolute values of the relative error; CCβ (mg L^{-1}), capability of detection with the probabilities of false positive and false negative fixed to 0.05.

Table 3 Principal component analysis of the variables RMSEC, RMSEC_CV, SEP, s_{yx}, b_0, b_1, ERROR and CCβ

Principal Component	Eigenvalue	Explained variance (%)	Accumulated variance (%)
1	3.943	49.29	49.29
2	2.317	28.96	78.25
3	0.773	9.66	87.91

Table 4 Global desirability (D) for each pretreatment (PRET)

PRET	D	PRET	D	PRET	D
141	0.0000	241	0.0000	341	0.0000
142	0.0000	242	0.0000	342	0.0000
143	0.2898	243	0.0000	343	0.0000
151	0.0000	251	0.0000	351	0.0000
152	0.0000	252	0.0000	352	0.0000
153	0.0000	253	0.0000	353	0.0000
161	0.2511	261	0.2521	361	0.2556
162	0.4269	262	0.3965	362	0.0000
163	0.9188	263	0.9105	363	0.8969
		521	0.0000	531	0.4284
		522	0.0000	532	0.4281
		523	0.6529	533	0.8486
		621	0.0000	631	0.4290
		622	0.9761	632	0.9497
		623	0.6618	633	0.8515

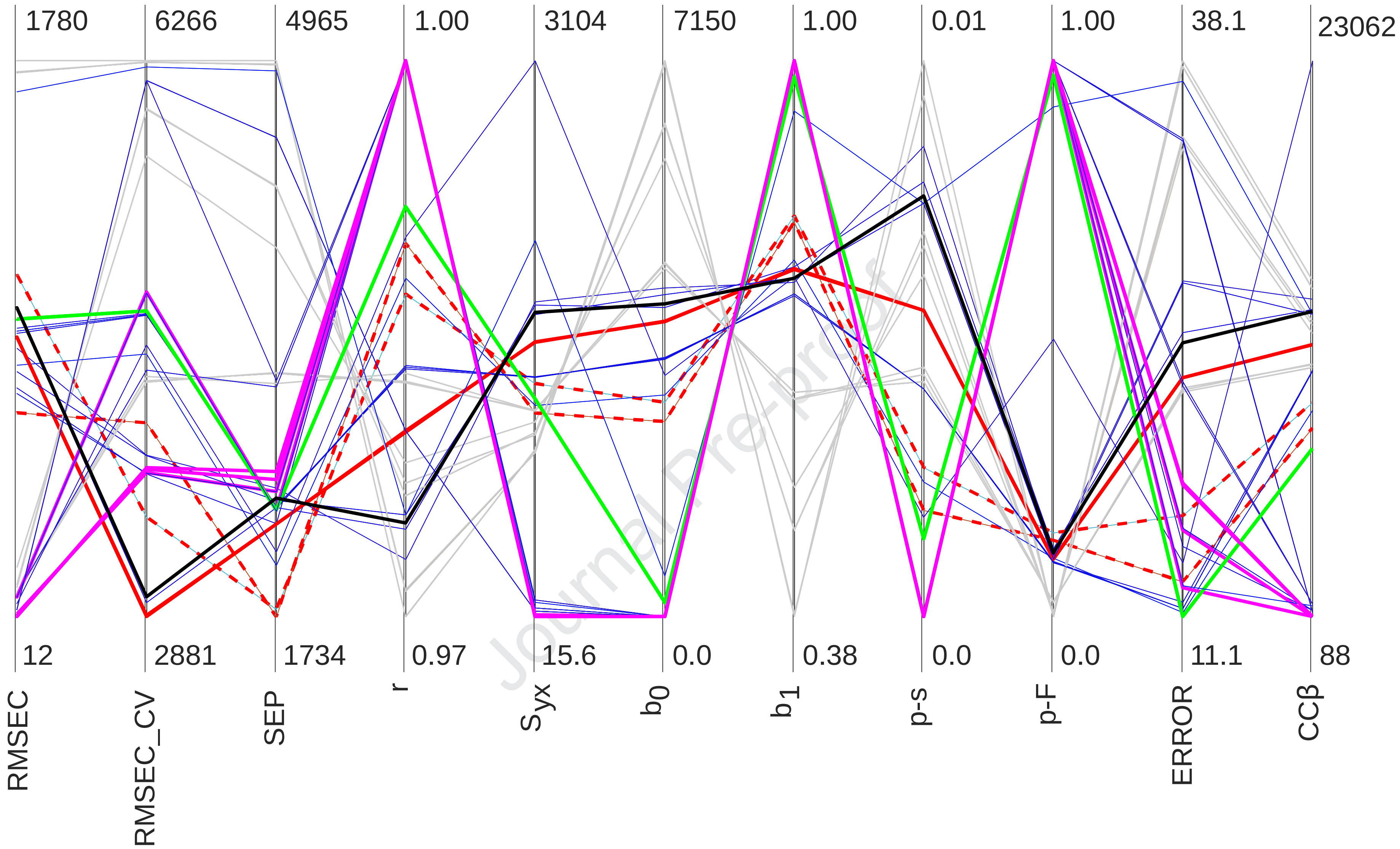
1780
6266
4965
1.00
3104
7150
1.00
0.01
1.00
38.1
23062
12
2881
1734
0.97
15.6
0.0
0.38
0.0
0.0
11.1
88
RMSEC
RMSEC_CV
SEP
r
S_{yx}
b_0
b_1
p-s
p-F
ERROR
CCβ

FIGURE 1

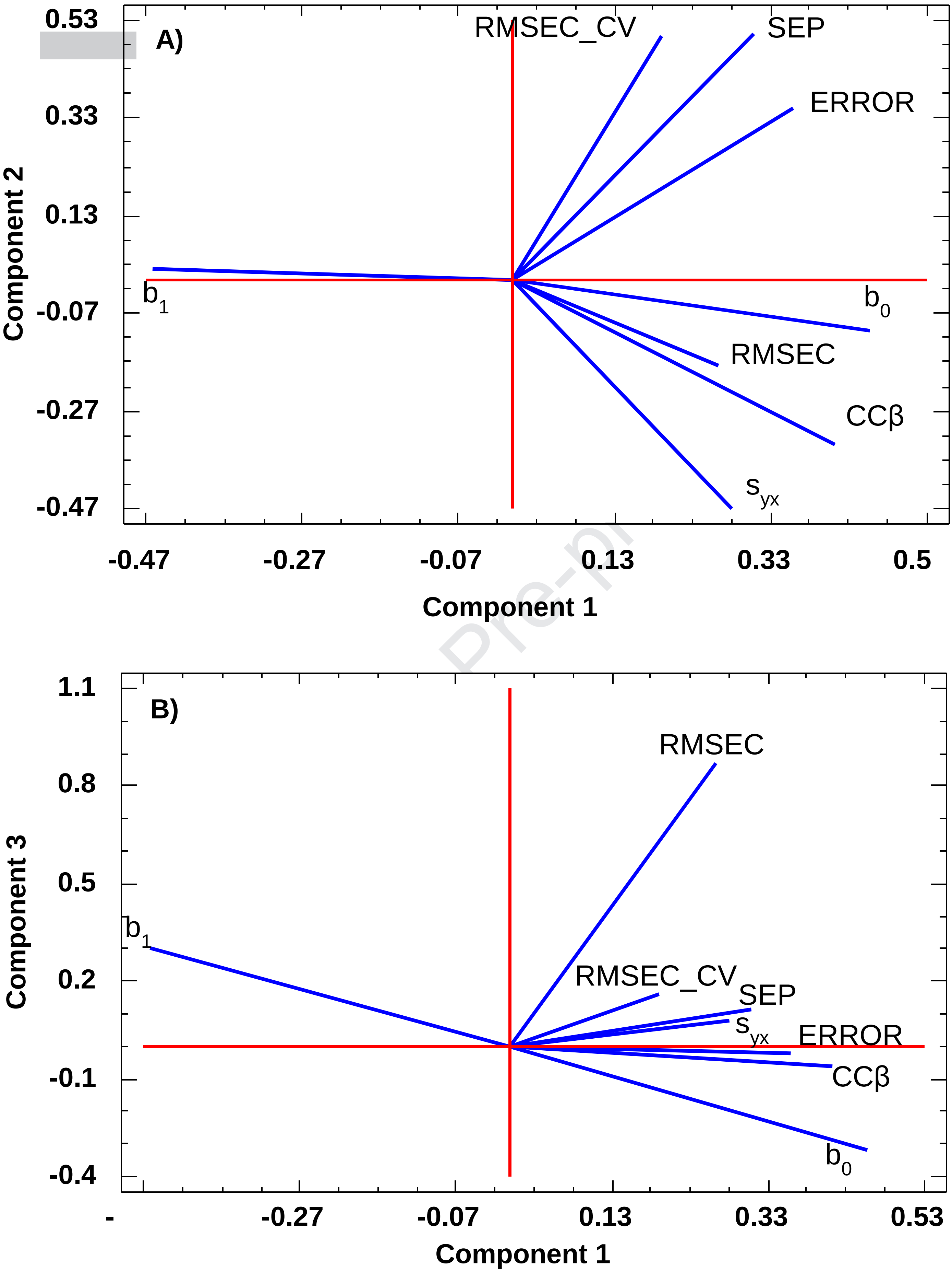
A)
Component 2
Component 1
0.53
0.33
0.13
-0.07
-0.27
-0.47
-0.47
-0.27
-0.07
0.13
0.33
0.5
RMSEC_CV
SEP
ERROR
b_1
b_0
RMSEC
CCβ
s_{yx}
B)
Component 3
Component 1
1.1
0.8
0.5
0.2
-0.1
-0.4
-
-0.27
-0.07
0.13
0.33
0.53
RMSEC
b_1
RMSEC_CV
SEP
s_{yx}
ERROR
CCβ
b_0

FIGURE 2

HIGHLIGHTS

Signal pretreatment mostly influences the vibrational spectroscopy methods

A procedure to obtain the best pretreatment is developed for PLSR calibration

The procedure is a multicriteria strategy that models the accuracy line quality

The performance of the procedure has been proved for calibration of BP3 by ATR-FTIR

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: