Abstract: Abstract
A simple and accurate analytical method for the simultaneous determination of three common dyes (Basic Blue 9, Brilliant Blue E-4BA, and Reactive Blue 2) in natural waters without prior separation of the solutes is proposed. A popular chemometric method, Partial Least Squares Regression PLS-1, was effectively applied for the resolution of the highly overlapping dyes. At the best modeling conditions, mean recoveries and relative standard deviations (R.S.D.) for dyes prediction by PLS-1 were found to be 102.1 (4.4), 95.7 (8.4), and 98.9 (6.2) for Basic Blue, Brilliant Blue, and Reactive Blue, respectively. Estimated limits of detection (L_D) using net-analyte signal concept were 0.11, 0.52, 0.49 mg L-1 for Basic Blue, Brilliant Blue, and Reactive Blue, respectively. The quantitative
determination of dyes spiked in real water samples was carried out successfully by the proposed PLS-1 method with satisfactory recoveries for dyes.
A Simple and Accurate Analytical Method for Determination of Three Commercial Dyes in Different Water Systems using Partial Least Squares Regression

Yahya S. Al-Degs\textsuperscript{a}, Amjad H. El-Sheikh\textsuperscript{a}, Ayman A. Issa\textsuperscript{a}, Mohammad A. Al-Ghouti\textsuperscript{b}, Mahmoud Sonjouk\textsuperscript{a}

\textsuperscript{a}Chemistry Department, The Hashemite University, P.O. Box 150459, Zarqa, Jordan.
\textsuperscript{b}Industrial Chemistry center, Royal Scientific Society, P.O.Box 1438, Amman, Jordan.

Abstract

A simple and accurate analytical method for the simultaneous determination of three common dyes (Basic Blue 9, Brilliant Blue E-4BA, and Reactive Blue 2) in natural waters without prior separation of the solutes is proposed. A popular chemometric method, Partial Least Squares Regression PLS-1, was effectively applied for the resolution of the highly overlapping dyes. At the best modeling conditions, mean recoveries and relative standard deviations (R.S.D.) for dyes prediction by PLS-1 were found to be 102.1 (4.4), 95.7 (8.4), and 98.9 (6.2) for Basic Blue, Brilliant Blue, and Reactive Blue, respectively. Estimated limits of detection (LOD) using net-analyte signal concept were 0.11, 0.52, 0.49 mg L\textsuperscript{-1} for Basic Blue, Brilliant Blue, and Reactive Blue, respectively. The quantitative determination of dyes spiked in real water samples was carried out successfully by the proposed PLS-1 method with satisfactory recoveries for dyes.

Keywords: Chemometry; Spectral overlap; PLS-1 calibration; Cationic and anionic dyes.

1. Introduction

Dyes (either natural or synthetic) can be classified into three types according to their chemical structure: cationic, nonionic, and anionic [1]. Both natural and synthetic dyes are heavily used by many industries including food, pharmaceutical, cosmetic, textile and leather industries [1]. The effluents of these industries are highly colored and disposal of these wastes into natural waters causes damage to the environment [1-2]. Direct, acid and reactive dyes are typical examples for anionic dyes [2-3]. It is difficult to remove dyes from effluents since they are sable to light, heat and oxidizing agents and are biologically non-degradable [2]. Some classes of dyes are harmful to aquatic life even at lower concentrations [2-3]. It is pointed out that less than 1.0 mg L\textsuperscript{-1} of dye content causes obvious water coloration [1]. Dye concentrations of 10 mg L\textsuperscript{-1} up to 25 mg L\textsuperscript{-1} have been cited as being present in dyehouse effluents [4]. After mixing with other water streams, the concentration of dyes is further diluted. The concentration of dyes could be in the μg/dm\textsuperscript{3} level in wastewater [5]. In response to concerns regarding the health risks associated with the use of reactive dyes, an increasing number of analytical methods have been developed in recent years for their determination. Several papers described the applications of new analytical...
procedures for the detection and determination of dyes in various matrices such as water, wastewater, urine, fish tissue, and animal feed [3,6]. The spectrophotometric determination of a mixture of dyes is a hard task in analytical chemistry due to the potential spectral interferences, which results in widely overlapped absorption bands [7]. For this determination, the conventional univariate calibration method is impractical, because of the contribution of one species to the absorption signals of other species, and vice versa. In the recent years, method development for resolution of highly overlapped spectra has increased exceptionally, because of the availability of powerful instrumentation and robust numerical methods. For example, derivative spectrophotometry [8], the H-point standard addition method [9], chemometric methods including classical least squares (CLS), principal-components regression (PCR), and partial least squares regression (PLS-1 or PLS-2) [10-12] have been frequently employed to overcome the problems of interference due to spectral overlapping. A limited number of studies have been reported on using multivariate calibration for analysis of multi-component mixtures of dyes [5,7,13]. While the use of spectrophotometric techniques is preferred because of their relatively low operation cost, their application would be limited due to their modest sensitivity [14]. In this work, the simultaneous determination of for Basic Blue, Brilliant Blue, and Reactive Blue in water was described. The quantification of dyes was developed by UV-Vis spectrophotometry in their mixtures by PLS-1. The figures of merit such as selectivity, sensitivity, limit of detection (LOD) and analytical sensitivity for the multivariate method were calculated. The final aim of this study is to quantify the dyes in different water systems with aid of optimized PLS-1 method.

2. Theoretical background of PLS-1 method

The following notations were adopted in this work, boldface capital letters for matrices, e.g., \( A \) and \( C \), superscript \( t \) for transposed matrices (or vectors), e.g., \( A^t \), boldface small characters for vectors, e.g., \( v \) and small characters for scalars, e.g., \( c \). \( \| \| \) stands for the Euclidian norm for the vector.

The Beer’s law model for \( m \) calibration standards containing \( l \) chemical components with spectra of \( n \) digitized absorbances can be presented in matrix notation as [10]:

\[
A = CK + E
\]  

(1)

Where \( A \) is the \( m \times n \) absorption matrix of calibration spectra, \( C \) is the \( m \times l \) concentration matrix of components (dyes in this study) concentrations, \( K \) is the \( l \times n \) matrix of absorptivity constants or simply the calibration matrix. \( E \) is the \( m \times n \) matrix of spectral errors that not fit by the model. The elements of \( K \)-matrix can be determined by measuring spectra of individual components; however, this dose not takes into accounts the interactions (or overlapping) between components. In literature, there are many chemometric techniques that was effectively solved Equation 1 and found the perfect relation between the absorbance and concentration. Among these calibration methods, Classical Least Squares (CLS), Principal Component Regression (PCR), Partial Least Squares, and Kalman Filter (KF) were the most employed.
methods in multivariate calibration [13]. Determination of dyes concentration in synthetic and real water samples using PLS-1 method was carried out according to the theories proposed in [10,13].

2.2. PLS-1 calibration method

Usually, multivariate calibration methods involve a calibration step in which the relationship between spectra and component concentrations is estimated from a set of calibration samples, and a prediction step in which the results of the calibration are used to predict the component concentrations in an unknown sample spectrum [10]. Partial least squares type 1 (PLS-1) is a factor analysis method which has many of the full spectrum advantages of its classical counterpart (CLS). The method keeps the advantages of inverse calibration, allowing us to carry out the analysis for one chemical component at a time [10]. In the PLS-1 type, all parameters are optimized for the determination of a single analyte of interest. In PLS-1, the calibration spectra can be represented as:

\[
A = TP^t + ER \tag{2}
\]

where \( A \) is an \( m \times n \) matrix containing the spectra of \( m \) calibration samples obtained at \( n \) wavelengths, \( P \) is a \( n \times h \) matrix containing the full spectrum vectors (loadings), \( T \) is an \( m \times h \) matrix of intensities (or scores) in the new coordinate system defined by the \( h \) loading vectors, and \( ER \) is the \( m \times n \) matrix of spectral residuals not fitted by the optimal PLS-1 model. The loading vectors contained in \( P \) are usually determined by an iterative algorithm, which also provides a set of orthogonal weight loading factors that form the \( n \times h \) matrix \( W \). The mutual relationship between \( A, P, T \) and \( W \) is expressed in the following equation:

\[
T = RW (P^tW)^{-1} \tag{3}
\]

In PLS-1, the matrix \( T \) is related to concentration by an inverse regression step as shown in equation 5:

\[
c_k = Tv + e_c \tag{4}
\]

where \( c_k \) is the \( m \times l \) vector of the concentrations of analyte \( k \) in the calibration samples, \( v \) is the \( h \times l \) vector of coefficients relating the scores to the concentrations and \( e_c \) collects the corresponding concentration residuals. In the prediction step, the spectrum \( r \) registered for an unknown sample is transformed into the sample score \( t_r \) by:

\[
t_r = (W^tP)^{-1}W^tr \tag{5}
\]

from which the concentration can be calculated as:

\[
c_{k,un} = t_r^tv \tag{6}
\]

where \( v \) is the vector of regression coefficient of Eq. (5).

2.2. Figures of merits for the analytical method

The selectivity, sensitivity and limit of determination can be calculated to study the quality of a given analytical method. Selectivity, can be expressed as [15]:

\[
SEL_k = ||s^*_k|| / ||s_k|| \tag{7}
\]
where $s_k^*$ is the projection of $s_k$ (the vector of the spectral sensitivities of compound $k$ in pure form) onto the so-called net analyte signal space [15]. Sensitivity for analyte $k$ can be calculated using the following equation [15]:

$$\text{SEN}_k = ||s_k^*||$$

(8)

Another important figure of merit is the limit of detection (LOD), which can be calculated using the following equation [16]:

$$\text{LOD}_k = 3 ||\varepsilon||/||s_k^*||$$

(9)

where $\varepsilon$ is a measure of the instrumental noise.

3. Experimental

3.1. Instrumentation and Software

The absorbance measurements were obtained using a double beam Unicam spectrophotometer (Cary 50 UV-Vis spectrophotometer). The UV-vis spectra were recorded over the wavelength range of 400-600 nm and digitized absorbance was sampled at 1.0 nm intervals and then transferred to a Pentium(IV) personal computer for subsequent analysis. The data treatment was carried out using MATLAB (version 7.0). The pH measurements were made with a WTW-Inolab (Germany) pH-meter using a companied glass electrode.

3.2. Reagents

Doubly distilled water and high purity reagents were used for all preparations of the standard and sample solutions. The selected dyes—which have a wide application in many industries—were; Basic Blue 9, Brilliant Blue E-4BA, and Reactive Blue 2. Dyes materials of analytical grade (> 99.9) were purchased from Aldrich Company. The chemical structures of dyes are illustrated in Fig. 1. The employed dyes were heavily used in clothes dyeing. Standard stock solutions ($1.00 \times 10^3$ mg L$^{-1}$) of each dye were prepared individually by dissolving 0.100 (±0.001) g in doubly distilled water in a 100.0 cm$^3$ volumetric flask. Dilute solutions were prepared by the appropriate dilution of the stock solution in doubly distilled water.

3.3. Optimization of dyes solution pH

Due to the potential effect of solution pH on dyes absorption intensities, the optimum pH (the one that ensure high sensitivity and selectivity for dyes analysis) should be selected beforehand. To achieve that purpose, both selectivity SEL and sensitivity SEN for each dye were calculated for each investigated pH. The net-analyte signal concept, originally developed by Lorber [15] was applied for sake of finding SEL and SEN (Eqs. 7-8 in section 2.2 above). The adjustment of pH of dyes solutions was carried out using 0.5 M HNO$_3$ or 0.5 M NaOH. The ionic strength of dye solutions was adjusted using pure NaCl. All used reagents used for pH and ionic strength adjustment were of high purity.
3.4. Development and validation of PLS-1 method

Due to the high spectral overlap (98% overlap) between dyes, large number of calibration samples were necessary to build PLS-1 model. Typical one-compound calibration experiments (univariate calibration) were carried out to establish the concentration ranges for the determination in the mixture. Solutions of absorbance values higher than 1.0 were diluted. Twenty eight ternary synthetic mixtures (maintained at pH 6.0 and ionic strength of 0.10 M NaCl) of dyes were carefully prepared. From these solutions, 20 solutions were selected as the calibration set and the rest of solutions were kept for validation. The calibration and prediction sets were presented as 3D plot so as to view the homogeneity between them (Fig. 2). To reduce the collinearity in the absorption matrix $A$, the design for calibration set was based on four-level fractional factorial designs according to Brereton’s procedure [17]. The prediction set was selected randomly. For best calibration results, the spectral region within the range (200-800 nm) was chosen. The number of experimental points ($\lambda$) per spectrum is 601 where the spectrum is subdivided into 1.0 nm intervals. Within this spectral region, maximum spectral information was available. Accordingly, the dimension of calibration absorption matrix ($A$) is 20×601, while the dimension for calibration concentration matrix ($C$) is 20×3. The concentration range of dyes in the prediction samples should be within the space of calibration samples which is clear from the high homogeneity of the 3D plot of the samples (Fig. 2).

3.5. Determination of dyes in real water systems

Natural water samples including sea, river, and treated wastewater were collected from different locations. Samples of water (1000 mL) were obtained directly form source and stored in polyethylene bottles at 10 °C. Prior spiking of natural samples with dyes, all samples were filtered through a cellulose membrane filter (Millipore) of 0.45 μm pore size to remove any suspended matter. Furthermore, the pH of water samples was adjusted to 6.0 using the diluted acid. The spectra of spiked solutions were recorded over the region 200-800 nm and the obtained data were digitized and analyzed by the proposed PLS-1 method to determine the concentration for dyes.

4. Result and discussion

Determination of dyes mixture in water or wastewater can be carried out using high performance liquid chromatography (HPLC), gas chromatography/mass spectrometry (GC-MS), liquid chromatography/mass spectrometry (LC-MS), and capillary electrophoresis (CE) [5]. However, chromatographic determination of dyes in the mixture is a lengthy procedure also previous clean-up and preconcentration steps are essential in some instances. Multivariate calibration methods (particularly PCR and PLS-1) have been applied to highly overlapped spectra or chromatograms [18-19]. The advantages of applying multivariate calibration methods were to minimize or eliminate sample preparation and also to avoid a preliminary separation step in complex matrices [19-20].
4.1. Extent of spectral overlap between dyes and the importance of chemometry

Figure 3 shows the visible absorption spectra of the three dyes used in this study. As can be seen, an important degree of spectral overlap occurs between studied dyes. The degree of spectral overlap was estimated as outlined in the literature [21]. The spectral overlap between Brilliant Blue and Reactive Blue was 92%. However, a lower overlaps were noted for Basic Blue with other two dyes (70–85%). Due to the significant spectral overlapping (85-92%), the conventional calibration procedures would have a limited application for quantitative determination. Therefore, the simultaneous determination of these dyes requires the application of PLS-1 method to overcome such a high spectral overlap.

The linear working concentration range of each dye was determined separately. These concentration ranges were useful in building both calibration and prediction sets. The linear concentration range of each dye was determined by regressing absorbance at the corresponding $\lambda_{\text{max}}$ against concentration at pH 6. Limit of detection (LOD) was taken as $3\sigma$/slope where $\sigma$ is the standard deviation of 10 measurements of the blank.

Three straight lines for dyes calibration with $r^2 = 0.9995$, 0.9993, and 0.9997 for Basic Blue, Brilliant Blue and Reactive Blue; respectively. Linearity was observed between 1.0 and 30.0 mg L$^{-1}$ for both Brilliant Blue ($\lambda_{\text{max}} = 605$ nm) and Reactive Blue ($\lambda_{\text{max}} = 615$ nm) and from 0.15 to 8.0 mg L$^{-1}$ for Basic Blue ($\lambda_{\text{max}} = 671$ nm). The LOD values were 0.048, 0.35, and 0.31 mg L$^{-1}$ for Basic Blue, Brilliant Blue and Reactive Blue; respectively. The molar absorptivity (in cm$^{-1}$ M$^{-1}$) for dyes were $6.1 \times 10^4$, $7.7 \times 10^3$, and $7.0 \times 10^3$ for Basic Blue, Brilliant Blue and Reactive Blue; respectively. Compare to the rest of dyes, Basic Blue has a high molar absorptivity value which will improve its analytic selectivity and sensitivity as will be shown later. The selected concentration ranges of dyes used in multi-component samples were selected in a way to avoid excessive absorbance of the mixtures [22].

4.2. Optimization of experimental conditions affecting dyes absorption

Many experimental conditions may affect the absorption characteristics of dyes, among these; pH, temperature and ionic strength of solution are the most important. The influence of pH on dyes absorption intensity was studied over a wide pH range (2-12) at constant solution ionic strength (0.1 M NaCl) and temperature 26 °C. The initial spectral studies indicated that the spectra of Reactive Blue and Brilliant Blue did not alter over the investigated pH range (2-12). However, large spectral changes were observed for Basic Blue especially in basic media (pH > 7.5). Accordingly, the pH value (or pH range) at which the maximum SEN and SEL for all dyes are achieved should be determined. As mentioned before, both SEN and SEL values (at each pH) were obtained using NAS concept and the results were all given in Table 1. As indicated from Table 1, Basic Blue has a high SEL and SEN values compare to the rest of dyes at pH ≤ 8. Above pH 8, a considerable decrease in the sensitivity is observed and this could be attributed to the changes the dye structure due to reactive with hydroxyl ions. Slight changes were observed in the values of SEL and SEN for Brilliant Blue and Reactive Blue dyes over the entire pH range. The high SEL and SEN values observed for Basic Blue could be attributed to the high molar absorptivity for this dye which is 10-
fold higher than the rest of dyes. Accordingly, the optimum pH range for dyes analysis is 2-8 and pH 6 was
selected from this range.

4.4. Calibration by PLS-1 and selection of the optimum number of latent variables (h)

Determination the optimum number of latent variables (h) to be used in PLS-1 method is essential to
achieve high prediction. This allows modeling of the system with the optimum amount of information,
avoiding overfitting [10]. To avoid overfitting, leave-one-out cross validation procedure was adopted [10].
Simply, for m calibration set (20 solutions in our case), the PLS-1 was performed on m-1 calibration
solutions, and use this calibration to predict the concentration in the sample left out. This process was
repeated 20 times until each sample has been left out once [10,18]. The selection of spectral region for
numerical analysis was carried by applying a moving window strategy to the calibration set itself as
outlined elsewhere [20]. The predicted concentration for each sample is then compared with the true
concentration value. PRESS (prediction error sum of squares), which measures the difference between
predicted concentration and true one, is then estimated for all calibration samples in the set. PRESS value
is also calculated after each increment in h as follows [18]:

\[
\text{PRESS} = \left( \sum_{i=1}^{m} (C_{i,\text{pred}} - C_{i,\text{act}})^2 \right)
\]

(11)

Where, m, C_{\text{pred}}, and C_{\text{act}} are the total number of calibration samples, predicted concentration and the actual
concentration of the dye respectively. The effectiveness of PLS-1 for dyes prediction in the validation set
was determined by calculating relative error of prediction (REP), root mean squares difference (RMSD),
standard error of calibration or prediction SEC(P), and square of the correlation coefficient (r^2).

Statistical parameters for PLS-1 method and the figures of merit for dyes determination are presented in
Table 2. As can be seen, the obtained latent variables were higher than 3 for Brilliant Blue and Reactive
Blue, indicating that the variability sources number in the presently studied system exceeds the number of
studied analytes (3 analytes). Figure 4 depicted the PRESS-Latent variable plot obtained from cross-
validation technique.

As indicated in Table 2, similar spectral regions were used in PLS-1 calibration for Brilliant Blue and
Reactive Blue. The perfect spectral region for Basic Blue was extended from 510 to 720 nm. High
prediction power was noted for PLS-1 for all dyes in the calibration samples as indicated from r^2 and
REP% values. Usually, the multivariate selectivity values ranges between zero (complete overlap between
analyte and other interferences) and unity (no or small overlap between analyte and other analytes or
interferences) [24]. The lower selectivities obtained for Brilliant Blue (0.432) and Reactive Blue (0.489)
compare to Basic Blue (0.872) were is expected due to the high spectral overlap between Brilliant Blue and
Reactive Blue as shown in Fig. 3. On the other hand, the high sensitivity value (0.0882) observed for Basic
Blue is expected due to the high molar absorptivity of this dye compare to the rest of dyes.
The multivariate LODs as estimated from NAS concept were 0.11, 0.52 and 0.49 for Basic Blue, Brilliant Blue, and Reactive Blue; respectively. The optimized PLS-1 calibration method was further validated by finding dyes contents in the validation set. The mean recovery percentages, R.S.D., SEP, and REP% obtained for prediction or validation set were presented in Table 3. The results were satisfactory indicating the successful application of the proposed method for simultaneous determination of the three dyes in water. Relatively speaking, the PLS-1 method was more effective in internal validation compared to external validation as indicated from the values of REP% for both sets. Comparison between SEC and SEP values allows identification of an overfitting or underfitting in the model, with more or less latent variables than necessarily required [13]. For each dye, the magnitude of SEC and SEP were fairly similar which confirmed the perfect selection of the number of latent variables in both cases.

4.5. Determination of dyes mixtures spiked in different real water samples using PLS-1 method

The next step was to test how well the proposed PLS-1 method will do when applied to dyes determination present in real water systems, considering the fact that the calibration set was designed using standards where other interferences were not accounted. As stated earlier, the main aim of this study was to develop an analytical method with a minimal sample preparation and find a simple analytical method to replace the current analytical methods [6]. Nine ternary mixtures with variable concentrations of dyes were spiked in different natural water samples. The concentration levels of added dyes were selected to be determined by PLS-1 method. The percentages recoveries and R.S.D. values obtained by PLS-1 were summarized in Table 4. The recoveries obtained were satisfactory in all the samples analyzed where the values of R.S.D. were less than 10%. The prediction power of PLS-1 can be considered acceptable taking into account the complexity of the sample being analyzed. The good agreement between the PLS-1 results and the spiked values is an indication of the effectiveness of the proposed method for simultaneous determination of Basic Blue, Brilliant Blue, and Reactive Blue in real samples. In fact, other interferences that present in natural water samples which where not considered in the calibration model, e.g., Cl\(^-\), SO\(_4\)\(^{2-}\), NO\(_3\)\(^-\), OM, etc., did not strongly interfere with dyes analysis. In fact, the spectral regions that selected for dyes analysis was within the visible region which probably the absorption of natural organic matters is not high in that region.
Conclusions

The results indicated that PLS-1 method is a rapid, easy and of low cost for simultaneous determination of three commercial dyes (two anionic dyes and one cationic) in different water systems. The proposed method could be used for the screening of dyes in water (e.g., in situ analyses) or as a quantification method in cases where the chromatographic ones cannot be implemented owing to cost limitations, lack of analytical instrumentation. Quantitative determination of dyes in real water samples was carried out successfully by the proposed method with recoveries ranging between 90–106 with R.S.D. values < 10%.

Acknowledgements

The financial support from the Hashemite university/deanship of academic research is gratefully acknowledged. Dr. Yahya would like to thank Mr. Khaled Akel for running the spectral analysis.

References


List of Figures:

**Fig. 1.** The chemical structures of dyes

**Fig. 2.** A 3D plot presentation of calibration and validation sets

**Fig. 2.** Absorption spectra of studied dyes

**Fig. 4.** PRESS values against number of latent variables
1) Reactive Blue 2 (anionic dye)

2) Brilliant Blue E-4BA (anionic dye)

3) Basic Blue 9 (cationic dye)

Fig. 1. Structural formulae of the studied dyes
Fig. 2. A 3D plot showing the composition of the three dyes used in PLS-1 calibration and validation.
Fig. 3. Absorption spectra of dye. Brilliant Blue 18 mg L$^{-1}$ (1), Reactive Blue 14 mg L$^{-1}$ (2) Basic Blue 1.0 mg L$^{-1}$. pH 6.0 and 0.10 M NaCl.
Fig. 4. Plot of PRESS against PLS-1 latent variables
List of Tables:

Table 1. SEL and SEN values

Table 2. Statistical parameters and the figures of merit for PLS-1 method

Table 3. Statistical parameters for external validation of PLS-1 method for dyes.

Table 4. Determination of dyes in different water systems by PLS-1 method
Table 1. The values of SEL and SEN obtained for dyes at different pH values

<table>
<thead>
<tr>
<th>pH</th>
<th>Basic Blue</th>
<th>Brilliant Blue</th>
<th>Reactive Blue</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SEL</td>
<td>SEN</td>
<td>SEL</td>
</tr>
<tr>
<td>2</td>
<td>0.822</td>
<td>0.0871</td>
<td>0.423</td>
</tr>
<tr>
<td>4</td>
<td>0.825</td>
<td>0.0873</td>
<td>0.444</td>
</tr>
<tr>
<td>6</td>
<td>0.872</td>
<td>0.0882</td>
<td>0.432</td>
</tr>
<tr>
<td>8</td>
<td>0.725</td>
<td>0.0752</td>
<td>0.452</td>
</tr>
<tr>
<td>10</td>
<td>0.720</td>
<td>0.0251</td>
<td>0.456</td>
</tr>
<tr>
<td>12</td>
<td>0.625</td>
<td>0.0110</td>
<td>0.462</td>
</tr>
</tbody>
</table>
Table 2. Statistical parameters and the figures of merit for PLS-1 method

<table>
<thead>
<tr>
<th>Calibration parameter</th>
<th>Basic Blue</th>
<th>Brilliant Blue</th>
<th>Reactive Blue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spectral region (nm)</td>
<td>510-720</td>
<td>440-680</td>
<td>430-705</td>
</tr>
<tr>
<td>Number of factors ($h$)</td>
<td>3</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>PRESS (mg^2 L^-2)</td>
<td>2.15</td>
<td>5.53</td>
<td>7.02</td>
</tr>
<tr>
<td>RMSD (mg L^-1) $^a$</td>
<td>0.36</td>
<td>0.53</td>
<td>0.58</td>
</tr>
<tr>
<td>REP% $^b$</td>
<td>1.1</td>
<td>2.9</td>
<td>4.0</td>
</tr>
<tr>
<td>SEC $^c$</td>
<td>0.15</td>
<td>0.24</td>
<td>0.13</td>
</tr>
<tr>
<td>$r^2$ $^d$</td>
<td>0.9986</td>
<td>0.9865</td>
<td>0.9803</td>
</tr>
<tr>
<td>SEN</td>
<td>0.0882</td>
<td>0.0223</td>
<td>0.0321</td>
</tr>
<tr>
<td>SEL</td>
<td>0.872</td>
<td>0.432</td>
<td>0.489</td>
</tr>
<tr>
<td>LOD$^e$ (mg L^-1)</td>
<td>0.11</td>
<td>0.52</td>
<td>0.49</td>
</tr>
</tbody>
</table>

a. RMSD value was calculated as [19]: $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} (C_{i,\text{pred}} - C_{i,\text{act}})^2$

b. REP% was calculated as [18]: $100 \times \left( \frac{\sum_{i=1}^{n} (C_{i,\text{pred}} - C_{i,\text{act}})^2}{\sum_{i=1}^{n} (C_{i,\text{act}})^2} \right)^{1/2}$, where $n$ is the number of samples in the prediction set (8 samples in this study).

c. SEC and SEP values were calculated as [22]: $SEC(P) = \sqrt{\frac{\sum_{i=1}^{n} (C_{i,\text{pred}} - C_{i,\text{act}})^2}{1 - n}}$

d. $r^2$ value was calculated as [19,23]: $r^2 = 1 - \frac{\sum_{i=1}^{n} (C_{i,\text{pred}} - C_{i,\text{act}})^2}{\sum_{i=1}^{n} (C_{i,\text{act}} - \bar{C})^2}$, where $\bar{C}$ is the average analyte concentration in the reference samples.

e. Assuming that $||\epsilon|| = 0.007$ absorbance units.
Table 3. Statistical parameters for external validation of PLS-1 method for dyes.

<table>
<thead>
<tr>
<th>No.</th>
<th>Added (mg L(^{-1}))</th>
<th>PLS-1 prediction (mg L(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Basic Blue</td>
<td>Brilliant Blue</td>
</tr>
<tr>
<td>1.</td>
<td>3.5</td>
<td>3.0</td>
</tr>
<tr>
<td>2.</td>
<td>3.5</td>
<td>1.5</td>
</tr>
<tr>
<td>3.</td>
<td>2.0</td>
<td>1.0</td>
</tr>
<tr>
<td>4.</td>
<td>1.0</td>
<td>1.5</td>
</tr>
<tr>
<td>5.</td>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>6.</td>
<td>3.5</td>
<td>3.0</td>
</tr>
<tr>
<td>7.</td>
<td>4.0</td>
<td>3.0</td>
</tr>
<tr>
<td>8.</td>
<td>3.5</td>
<td>3.5</td>
</tr>
</tbody>
</table>

*Statistical analysis*

<table>
<thead>
<tr>
<th></th>
<th>Average recovery±R.S.D. (n = 8)</th>
<th>SEP</th>
<th>REP%</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEP</td>
<td>102.1 (4.4)</td>
<td>0.14</td>
<td>1.6</td>
</tr>
<tr>
<td>REP%</td>
<td>95.7 (8.4)</td>
<td>0.22</td>
<td>3.1</td>
</tr>
<tr>
<td>SEP%</td>
<td>98.9 (6.2)</td>
<td>0.14</td>
<td>5.1</td>
</tr>
<tr>
<td>Water body no.</td>
<td>Water body</td>
<td>Mixture</td>
<td>Added</td>
</tr>
<tr>
<td>---------------</td>
<td>-----------------------</td>
<td>-------------</td>
<td>-------</td>
</tr>
<tr>
<td>River water</td>
<td>1</td>
<td>Basic Blue</td>
<td>4.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Brilliant Blue</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reactive Blue</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Basic Blue</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Brilliant Blue</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reactive Blue</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Basic Blue</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Brilliant Blue</td>
<td>4.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reactive Blue</td>
<td>4.5</td>
</tr>
<tr>
<td>Sea water</td>
<td>4</td>
<td>Basic Blue</td>
<td>4.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Brilliant Blue</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reactive Blue</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>Basic Blue</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Brilliant Blue</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reactive Blue</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>Basic Blue</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Brilliant Blue</td>
<td>4.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reactive Blue</td>
<td>4.5</td>
</tr>
<tr>
<td>Textile treated wastewater</td>
<td>7</td>
<td>Basic Blue</td>
<td>4.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Brilliant Blue</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reactive Blue</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Basic Blue</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Brilliant Blue</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reactive Blue</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>Basic Blue</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Brilliant Blue</td>
<td>4.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reactive Blue</td>
<td>4.5</td>
</tr>
</tbody>
</table>

1. The employed river water contains: $[\text{Cl}^-] = 0.12 \text{ M}$, total hardness (as $\text{CaCO}_3$) = 680 mg L$^{-1}$, and alkalinity (as $\text{CaCO}_3$) = 75 mg L$^{-1}$.
2. The employed sea water contains: $[\text{Cl}^-] = 0.45 \text{ M}$, total hardness (as $\text{CaCO}_3$) = 850 mg L$^{-1}$, alkalinity (as $\text{CaCO}_3$) = 220 mg L$^{-1}$.
3. The employed textile treated wastewater contains: total hardness (as $\text{CaCO}_3$) = 550 mg L$^{-1}$, $[\text{SO}_4^{2-}] = 110$ mg L$^{-1}$, $[\text{NO}_3^-] = 45$ mg L$^{-1}$, BOD = 13 mg L$^{-1}$, COD = 20 mg L$^{-1}$.
4. Expressed in mg L$^{-1}$.
5. Relative standard deviation ($n = 3$).