CHROMATIC ANALYSIS OF ACID DYES BY UV-Vis DIFFUSE REFLECTANCE SPECTROSCOPY AND CHEMOMETRICS

Perla M. Gelacio-Valdéz, Miguel Velázquez-Manzanares, Judith Amador-Hernández (Corresponding author)

Department of Analytical Chemistry, Faculty of Chemical Sciences, Autonomous University of Coahuila

CP 25280, Saltillo, Coahuila, Mexico

Tel (52) 844 415 5392 E-mail: judith.amador@uadec.edu.mx

The research was financed by CONACyT Mexico (COAH-2013-C24-206199).

Abstract

This work explores the influence of pH on the chromatic coordinates in the CIELab color space of four synthetic dyes widely used in the wool textile industry: Blue 2R, Yellow 4GN, Red 2B, and Bordeaux B (CAS no. 70209-96-0, 72479-28-8, 72017-66-4, and 52333-30-9, respectively). Due to the presence of at least one sulfonic group in the chemical structures, the chromatic coordinates change as the acidity of the aqueous medium increases in all cases. For Blue 2R and Bordeaux B in solution, chromatic coordinates in CIELab were predicted by Partial Least Squares Regression from absorption spectra and pH. Variations between Partial Least Squares Regression-estimated and CIELab color space coordinates have led to errors of less than 1 % in all cases. On the other hand, Soft Independent Modeling by Class Analogy was used to differentiate the dyes in aqueous media from absorption spectra regardless of pH. Thus, it is demonstrated the predictive capacity offered by multivariate mathematical tools such as Partial Least Squares Regression or Soft Independent Modeling by Class Analogy. However, there is still much to explore in the binomial chromatic coordinates-chemometrics.

1. Introduction

Color is a perception that results from a visual stimulation of the environment, the interpretation of which is subjective. In humans, the perceived color depends on the individual and physiological or anatomical traits, culture, age, sex, and mood, among other factors (Calvin et al, 2013; O'Connor, 2015). Therefore, it is crucial to characterize color through an instrument and employ a numerical system; colorimetry does it. Chromatic coordinates depend on the source of illumination, the object, and the observer.

Additionally, color spaces are systems that allow expressing the color of an object using a numerical notation based on three perceptual descriptors or attributes

(Capilla et al, 2002; Holtzschue, 2017):

a) Hue. It represents the predominant wavelengths that capture the human eyes reflected by an object so that it can be affirmed if it is red, blue, or yellow, for example.

b) Luminosity or value. It defines the amount of light reflected by an illuminated object such that light color is appreciated when the surface reflects a large amount of light or very dark when it reflects almost nothing.

c) Saturation. It corresponds to the intensity of the color in the transition from vivid to muted colors.

The CIELab color space (also called CIELAB) is widely used in almost all fields. It was proposed by the CIE (*Commission Internationale de l'Eclairage*) in 1976 to reduce one of the main problems of the original *Yxy* space: that equal distances in the chromaticity diagram *x,y* did not correspond to perceived equal differences in color. In particular, L* indicates luminosity in this space, so it varies from black to perfect white; a* and b* are the chromaticity coordinates. On a* axis, negative values correspond to a green hue while positive values are in the scope of red. Regarding b*, positive values correspond to the yellow hue and the negative ones to the blue. In all cases, the scales range from 0 to 100 (Ohta and Robertson, 2006). On the other hand, a color difference exists when the sample has color coordinates other than the standard or reference sample in a tolerance range; color differences are expressed as ΔL*, Δa*, and Δb*, respectively. Conventionally, unique instrumentation and software are needed to estimate sample chromatic coordinates.

In another context, pattern recognition is a branch of chemometrics that identify trends between data for classification or discrimination purposes. Principal Component Analysis (PCA) is an unsupervised pattern recognition technique based on data compression, where samples will be represented through new variables or principal components (PC), which are linear combinations of the original variables, according to the expression:

$$t_1 = w_{11}x_1 + w_{12}x_2 + w_{13}x_3 + \ldots + w_{1n}x_n \quad (1)$$

where t_1 is the first principal component, from x_1 to x_n are the original variables, and from w_1 to w_n are regression coefficients (Brereton, 2007).

Soft Independent Modeling of Class Analogies (SIMCA) is a supervised pattern recognition technique based on PCA that starts from a series of training samples represented by *n* variables, in which it is known *a priori* how many groups there are and which samples correspond to which groups. The discrimination model is established from these training samples, which will allow recognizing later to which group an unknown sample is assigned or if it does not correspond to any of the groups under study (Miller and Miller, 2005). Thus, Melendez et al (1999) used SIMCA and Genetic Inside Neural Network to characterize clarets and rose wines from Rioja using CIELAB parameters and spectrophotometric data. Also, Alonso-Salces et al (2004) used SIMCA to confirm the authenticity of ciders by determining total polyphenol contents and CIELAB chromatic parameters in Basque and French ciders.

On the other hand, Partial Least Squares Regression (PLS) is a multivariate calibration technique that allows the prediction of one or more quantitative properties of unknown samples. First, the PCA-based algorithm of PLS analyzes a set of

calibration samples represented by multiple independent variables and known values of the properties to be predicted; later, the mathematical model is validated from the second series of known samples. Different authors have successfully applied PLS to the chromatic analysis in CIELAB and distinct spectroscopic techniques. For example, Feng et al (2020) proposed the non-destructively determination of pH and solid soluble content in red bayberry by PLS and Least Square-Support Vector Machine algorithms combined with six different color spaces, CIELAB between them. Also, Yergenson and Aston (2020) evaluated the use of online near IR spectroscopy, the L* value from the CIELAB color space in the visible region, and PLS to determine coffee roast degree toward controlling acidity. Additionally, Lillotte et al (2021) used PLS to analyze UV-visible spectra and chromatic coordinates to study the density of compressed materials to develop a process analytical technology for quality control.

This work studied the influence of pH in aqueous media on the chromatic coordinates in CIELAB of four synthetic dyes for coloring wool by UV-vis Diffuse Reflectance Spectroscopy. Also, PLS was used to predict chromatic coordinates for two dyes in CIELAB based on the absorption spectra and pH. Finally, SIMCA was applied to discriminate between the dyes of interest, based on their absorption spectra and regardless of pH.

2. Experimental

2.1 Instrumentation

A UV-Visible spectrophotometer (Perkin Elmer, model Cary 300) was used, equipped with a diffuse reflectance sphere accessory (DRA-CA-30I), a calibrated Spectralon Reflectance Standard, and a transmittance cuvette holder. Data treatment was developed with the software WinUV Color v.4 (Agilent Technologies), Origin Pro 8 (OriginLab Corporation), and Pirouette v.4.2 (Infometrix).

2.2 Reagents and materials

All reagents were at least analytical grade. The dyes were Lanaset® from Huntsman Mexico: a) Blue 2R, CAS 70209-96-0; b) Yellow 4GN, CAS 72479-28-8; c) Red 2B, CAS 72017-66-4; d) Bordeaux B, CAS 52333-30-9. For each analyte, 200 mg L^{-1} aqueous solutions were prepared and stored at 8°C in amber glass bottles for one month. Tri distilled water was from Jalmek.

2.3 Procedure

For acid-base titrations, a 250 mL beaker with constant stirring was available, into which a pH electrode and thermometer were inserted, and the dye solution in NaCl 0.1 M was added. Then, adequate volumes of HCl or NaOH solutions at variable concentrations were added, registering increases in pH of 0.2 units, approximately. In each pH condition, the transmittance spectrum was recorded by UV-Vis Diffuse Reflectance Spectroscopy, with a resolution of 1 nm in the spectral interval from 380 to 750 nm.

2.4 Data treatment

Chromatic coordinates L*, a*, and b* of the CIELAB color space were estimated from transmission spectra by the WinUV Color software, using the observer at 10° and the illuminant D65. Also, transmission spectra were transformed into absorption spectra by the same software. Then, absorption spectra of each dye recorded at each pH condition were organized in matrix arrangements for chemometric analysis. Centering on the mean was applied as data pretreatment, and the optimal number of PC was estimated for each data series, considering the standard calibration error and cumulative variance. Finally, calibration PLS-based and discrimination SIMCA-based models were constructed and validated. For PLS, Table 1 describes the calibration and validation sample sets for Blue 2R and Bordeaux B, while for SIMCA, the training and validation sample sets for all studied dyes are described in Table 2. In both cases, the chromatic coordinates in CIELAB and pH were included.

Tables 1 and 2

3. Results and discussion

Chemical structures of dyes of interest are presented in Figure 1. As can be recognized, all four compounds share the following characteristics:

1. They absorb radiation in the visible region of the electromagnetic spectrum since they exhibit $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ transitions.
2. Have numerous conjugated chromophore groups for electronic resonance.
3. Exhibit auxochrome groups, specifically sulfonic acids and amines.

Although the presence of auxochrome groups is not directly responsible for their color, they modify the tones of each compound, in addition to turning them into water-soluble substances. On the other hand, these colorants have been designed for the dyeing of natural textile fibers; for this, they incorporate functional groups into their structures that turn them into acids and guarantee their complete dissociation in water since they will bind to their substrates with positive charges through electrostatic interactions. Table 3 shows the relationship between the radiation absorbed in the electromagnetic spectrum by a chemical structure and the observed color for that substance (Bujdud, 1999). In this work, transmittance spectra were transformed to the absorption mode since the last are the conventional spectra recognized for colored substances and can be registered in conventional UV-vis spectrophotometers.

Figure 1

Table 3

3.1 Influence of pH on the CIELab chromatic coordinates of acid dyes

3.1.1 Yellow 4GN

Figure 2a shows the UV-Vis absorption spectra of the 4GN yellow dye in aqueous media by modifying the pH of the medium between 4.8 to 0.4. As can be recognized, the spectrum exhibits an absorption maximum at 402 nm, which shows a hypochromic effect as the pH decreases. This tendency was reversible, from acid to neutral conditions. The absorption spectra remain constant in the pH range of 1.5 to

4.8 since the A^- species (conjugate base) predominates. Below pH 1.5, absorbance decreases due to the pH being in pKa ± 2, where the acidic and basic species coexist simultaneously. However, a second absorption maximum does not appear in the series. On the other hand, as there is no appearance of a second band, the yellow hue remains constant with a slight tendency to green (Figure 2b), recognizing the most significant variation in the saturation of the predominant coordinate (+b*), which decreases with pH. The chromatic coordinate -a* also decreases, but to a lesser extent, while the brightness (L*) practically remains constant.

Figure 2

3.1.2 Blue 2R

It is presented the absorption spectra of the dye by modifying the pH between 0.0 and 5.3 in Figure 3a. As can be seen, the spectra present an absorption band with a maximum at 590 nm, in which there is a hypochromic effect as the pH decreases (Figure 3a). The observed effect was reversible, going from a highly acidic to a neutral medium. An isosbestic point was observed at 520 nm, and the second band of low intensity with a maximum of 443 nm. The absorption maximum at 590 nm remains constant above pH=3.5 due to the only presence of the basic form of the dye; less evident is the tendency for the maximum at 443 nm due to lower absorbance signals than those for 590 nm.

Furthermore, the appearance of a second band with hypsochromic displacement when the pH decreases has a representative effect on the three chromatic coordinates of the dye (Figure 3b). The chromaticity changes as the pH decrease, from a low saturated blue hue with a slight tendency to green to a barely perceptible blue with a barely perceptible tendency to red (very close to the achromatic point). The change in saturation of the blue hue is significant since $\Delta b^* = 23$, while a* changes from the green axis (-9.9) to the red (+2.1); brightness increases by eight units. In other words, the color of the solution changes significantly in the interval pKa ± 2 due to the presence in different proportions of two chemical species, the acid and the conjugate base (pH between 0 and 3.5). On the other hand, the decrease in pH leads to spectra of lower intensity throughout the spectral interval, which produce a significant decrease in saturation of the hue, remaining close to the achromatic axis ($a^* = 0$, $b^* = 0$, regardless of L*).

Figure 3

3.1.3 Red 2B

Although there is a highly conjugated sulfonic acid in the chemical structure of Red 2B, the presence of chromium bonded to atoms with free pairs of electrons (N and O) is recognized so that the electronic transitions $n \rightarrow \pi^*$ play a significant role in the absorption of radiation by this compound. Figure 4a shows the absorption spectra of the dye in aqueous media between 0.5 y 5.6. As the pH decreases, a single absorption band at 480 nm is recognized with a hypochromic effect.

Regarding the chromatic coordinates of the solution in the CIELAB space (Figure 4b), the red hue with a tendency to yellow remains constant concerning pH in acidic conditions since there is no appearance of a second band nor displacement of the maximum, so the chromaticity is in the orange zone. The most significant

variation is recognized in the predominant coordinate +b*, which decreases by 12.7 units as the pH decreases. On the other hand, the chromatic coordinate +a* varies only three units while the luminosity remains practically constant. In summary, the decrease in the pH of the medium diminishes the saturation while keeping chromaticity and luminosity constant.

Figure 4

3.1.4 Bordeaux B

It is the compound with the highest molecular weight of the series of dyes under study, which is soluble in water due to the presence of two sulfonic acids in its structure and not just one as in the other compounds. Figure 5a shows the absorption spectra of the compound in aqueous media by modifying the pH with the addition of HCl. As can be seen, the spectra show an absorption maximum at 520 nm, while a hypochromic effect is observed as the pH decreases due to the sulfonic group being protonated (remains neutral), and its negative charge cannot enter into resonance with the rest of the conjugated bonds. From acid-base equilibrium, at pH > 2, the A^- species (conjugate base) predominates; in the pH range pKa ± 2, the acidic and basic species coexist simultaneously.

Regarding its chromatic coordinates in the CIELAB space (Figure 5b), the predominant coordinate for this compound is +a*, so the hue or tone is red with low saturation and a barely perceptible tendency to blue. As the pH decreases over the range of +a* from +41.6 to +26.0, the complementary chromatic coordinate, -b*, decreases by just two units, and the luminosity or L* increases slightly. In general, as pH decreases, the saturation decline, and the luminosity increases, but without a change in hue.

Figure 5

3.2 PLS for multivariate calibration of chromatic coordinates

Blue 2R and Bordeaux B were the dyes that showed significant color variations as a function of pH. Therefore, they were considered for PLS analyses to predict CIELab color coordinates in acid media. Table 1 shows calibration and validation data sets for both dyes.

The optimal number of PC to build calibration models was selected from the local minimum of cumulative variance and standard calibration error criteria from calibration sets; one (Blue 2R) and two (Bordeaux B) PC were chosen since low standard calibration errors were observed and cumulative variances higher than 99 % were obtained in both cases.

For Blue 2R, Table 4 shows the chromatic coordinates predicted in the validation series consisting of 7 samples (not considered in the calibration); the calibration model has been built with one PC for each chromatic coordinate. In all cases, the absolute color differences between the mathematically predicted value and the experimental value are less than 1, which is satisfactory in terms of tolerance to color differences on a scale of 0 to 100 for each chromatic coordinate. Additionally, Table 5 integrates the chromatic coordinates predicted for Bordeaux B in the validation series; as can be seen, color differences between the mathematically

predicted and the experimental value is also less than one unit, which shows the capacity of the model for color prediction as satisfactory. Finally, the total color difference in CIELab (ΔE*ab) was calculated according to:

$$\Delta E^*ab = [(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2]^{1/2} \quad (2)$$

where ΔL*, Δa*, and Δb* are color differences between predicted and measured values in each chromatic coordinate. For both colors, ΔE*ab values were below one unit, corresponding to differences imperceptible to the human eye.

3.3 SIMCA for classification of samples according to their chromatic coordinates

Classification or discrimination between samples is becoming increasingly important in industry, especially if the process can be automated based on instrumental data. In this case, SIMCA was used to recognize the dye in solution from its absorption spectrum and independently of pH. The four-training series shown in Table 2 were integrated into a single data set, where the absorption spectra at different pH conditions were related to the corresponding dye. The model of discrimination was based on the Euclidean distance between samples derived from PCA; to obtain cumulative variances higher than 99 %, the number of PC selected to construct the training set were: a) one for Yellow 4GN, b) two for Bordeaux B, c) two for Blue 2R, and d) one for Red 2B.

Tables 6 and 7 present the classification results obtained by SIMCA for training and validation sets, respectively, by considering a binary response (true/false) in dye recognition. As can be seen, true positive (TP) results prevail in all cases. That is, if in the sample there was a particular colorant, applying the SIMCA model to the absorption spectrum allows identifying which dye it is, even when the pH of the medium changes and with it, the chromatic coordinates of the dissolution. For all cases, sensitivity was estimated according to:

$$Sensitivity\ (\%) = TP*100/Total\ number\ of\ samples \quad (3)$$

where the sensitivity is the ability of the estimator to give a satisfactory classification of samples. Sensitivities of 98 % (training series) and 96 % (validation series) were found, which correspond to good values.

Also, the discriminant capacity of SIMCA in both sample series is readily appreciated by observing the projection of groups from both the training series (Figure 6a) and the prediction series (Figure 6b), considering the vector space given by the first two principal components. Remember that in these systems, the greater the Euclidean distance between samples, the more significant differences between them. Thus, having areas of each quadrant associated with each dye, the model has an excellent discriminate capacity to recognize what coloring is treated.

4. Conclusions

UV-Vis diffuse reflectance spectroscopy provided spectral and chromatic information from acid dyes in solution to relate them to studying color variation as a function of pH. For all dyes, the chromatic coordinates in the CIELab system changed in acidic conditions, mainly saturation; for solutions of dyes with acidic or basic groups in their chemical structure, pH is a factor that can be manipulated to reach differences in saturation and even chromaticity.

Partial Least Squares Regression was satisfactorily used with the colorants Blue 2R and Bordeaux B to predict their three CIELab chromatic coordinates (a*, b*, L*), taking as input information the UV-Vis absorption spectra and pH. In all cases, variations between estimated vs. actual coordinates (Δa*, Δb*, ΔL*) and ΔE*ab were less than one unit on a scale of 1 to 100, which correspond to differences imperceptible to the human eye. In this way, it is possible to determine the color of a solution with a conventional spectrophotometer and chemometric software instead of using more sophisticated equipment or specialized software for color characterization, which increases the cost.

On the other hand, it was possible to identify the presence of one of the four dyes in the medium, based only on their absorption spectra, independently of the pH, through Soft Independent Modeling by Class Analogy. Furthermore, the discrimination model allowed to recognize colorants correctly in solution in more than 95% of cases since sensitivities of 98 and 96 % for training validation sets were estimated, respectively.

In this way, it is demonstrated that chemometric tools with great prediction capacity, such as Partial Least Squares Regression and Soft Independent Modeling by Class Analogy, are helpful and exhibit significant advantages such as the reduction of cost of this type of study. It is necessary to explore new chemometric applications in the characterization of color through numerical systems such as CIELAB color space since this combination has been little explored and offers exciting advantages.

References

Alonso-Salces, R. M., Guyot, S., Herrero, C., Berrueta, L. A., Drilleau, J.-F., Gallo, B., & Vicente, F. (2005). Chemometric classification of Basque and French ciders based on their total polyphenol contents and CIELab parameters. *Food Chemistry*, 91, 91-98.

Brereton, R. (2007). Applied chemometrics for scientists. Wiley, Chichester UK.

Bujdud, J. M. (1999). Practical color measurement. Thesis, Master of Science (Optics). University of Guanajuato, Mexico.

Capilla, P., Artigas, J., & Pujol, J. (2002). Fundamentos de colorimetría. Universitat de València, Valencia.

Feng, J., Jiang, L., Zhang, J., Zheng, H., Sun, Y., Chen, S., Yu, M., Hu, W., Shi, D., Sun, X., et al (2020). Nondestructive determination of soluble solids content and pH in red bayberry (*Myrica rubra*) based on color space. *Journal of Food Science and Technology*, 57, 4541-4550.

Holtzschue, L. (2017). Understanding color: an introduction for designers, fifth ed. Wiley, New York.

Lillotte, T. D., Joester, M., Frindt, B., Berghaus, A., Lammens, R. F., & Wagner, K. G. (2021). UV-Vis spectra as potential process analytical technology (PAT) for measuring the density of compressed materials: evaluation of the CIELAB color space. *International Journal of Pharmaceutics*, 603, 120668.

Melendez, M. E., Ortiz, M. C., Sanchez, M. S., Sarabia, L. A., & Iniguez, M. (1999). Chemometric characterization of the claretes and rose wines of the certified denomination of origin Rioja using CieLab parameters. *Quimica Analitica*, 18, 119-126.

Miller, J. N., & Miller, J. C. (2005). Statistics and chemometrics for analytical chemistry, fifth ed. Pearson Education Limited, Essex UK.

Ohta, N., & Robertson, A. (2006). Colorimetry: fundamentals and applications. Wiley, Chichester UK.

Or, C. K. L., & Wang, H. H. L. (2014). Color–concept associations: a cross-occupational and cultural study and comparison. *COLOR Research and Application*, 39, 630-635.

Yergenson, N., & Aston, D. E. (2020). Online determination of coffee roast degree toward controlling acidity. *Journal of Near Infrared Spectroscopy*, 28, 175-185.

O'Connor, Z. (2015). Colour, contrast and gestalt theories of perception: the impact in contemporary visual communications design. *COLOR Research and Application*, 40, 85-92.

Tables

Table 1. Composition of sample sets for PLS analysis.

Dye	Step	Number of samples	Interval studied			
			pH	a*	b*	L*
Blue 2R	Calibration	15	0.02 to 5.26	+2.1 to – 9.9	-3.5 to -26.3	89.2 to 81.0
	Validation	7	0.20 to 4.97	-0.04 to -10.0	-7.2 to -27.0	87.4 to 81.2
Bordeaux B	Calibration	12	0.30 to 5.75	+26.0 to +41.6	-2.2 to -4.1	78.1 to 72.4
	Validation	7	1.27 to 4.91	+36.2 to +42.4	-2.4 to -3.9	75.5 to 72.8

Table 2. Composition of sample sets for SIMCA analysis.

Dye	Step	Number of samples	Interval studied			
			pH	a*	b*	L*
Yellow 4GN	Training	12	0.40 to 4.83	-14.4 to -17.2	+43.6 to +57.8	98.9 to 98.2
	Validation	6	0.61 to 4.36	-15.0 to -17.1	+46.1 to +57.9	98.9 to 98.2
Bordeaux B	Training	12	0.30 to 5.75	+26.0 to +41.6	-2.2 to -4.1	78.1 to 72.4
	Validation	6	0.60 to 5.20	+30.3 to +42.4	-2.5 to -4.0	77.8 to 72.6
Blue 2R	Training	15	0.02 to 5.26	+2.1 to -9.9	-3.5 to -26.3	89.2 to 81.0
	Validation	7	0.10 to 4.75	+0.9 to -9.9	-5.3 to -27.1	88.3 to 81.0
Red 2B	Training	12	0.50 to 5.64	+21.5 to +24.8	+39.4 to +50.7	88.4 to 86.7
	Validation	6	0.80 to 5.37	+23.6 to +24.8	+45.8 to +50.7	87.6 to 86.7

Table 3. Relationship between radiation absorbed by substances and observed color.

Absorbed wavelength (nm)	Observed color
400-435	Yellow-green
435-480	Yellow
480-490	Orange
490-500	Red
500-560	Purple
560-580	Violet
580-595	Blue
595-605	Green-blue
605-700	Blue-green

Table 4. Color coordinates that were predicted in the Blue 2R validation series.

pH	a* predicted	|Δa*|	b* predicted	|Δb*|	L* predicted	|ΔL*|	ΔE*ab
4.97	-9.95	0.05	-27.42	0.42	81.12	0.08	0.43
4.01	-9.99	0.01	-27.49	0.49	81.10	0.00	0.49
3.12	-9.65	0.05	-26.82	0.22	81.32	0.08	0.24
2.52	-9.23	0.03	-26.00	0.20	81.58	0.08	0.22
2.05	-8.63	0.03	-24.81	0.0	81.98	0.02	0.04
0.96	-6.26	0.26	-20.15	0.85	83.50	0.00	0.89
0.20	+0.03	0.07	-7.73	0.53	87.56	0.16	0.56

Table 5. Color coordinates that were predicted in the Bordeaux B validation series.

pH	a* predicted	|Δa*|	b* predicted	|Δb*|	L* predicted	|ΔL*|	ΔE*ab
4.91	42.40	0.03	-3.90	0.04	72.80	0.05	0.07
4.29	42.50	0.01	-3.80	0.0	72.90	0.07	0.07
3.71	42.30	0.09	-3.80	0.17	73.10	0.02	0.19
3.09	42.2	0.01	-3.70	0.10	73.10	0.00	0.10
2.88	41.30	0.19	-3.60	0.26	73.70	0.02	0.32
2.48	39.50	0.25	-3.30	0.15	74.30	0.10	0.31
1.27	36.20	0.05	-2.40	0.41	75.50	0.05	0.42

Table 6. 2x2 Contingency table for classification of the training set samples by SIMCA.

		Predicted dye				
		Yellow 4GN	Bordeaux B	Blue 2R	Red 2B	Without classification
Real dye	Yellow 4GN	12	0	0	0	0
	Bordeaux B	0	12	0	0	0
	Blue 2R	0	0	14	0	1
	Red 2B	0	0	0	12	0

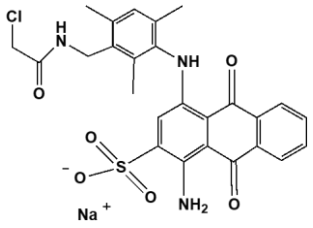
Table 7. 2x2 Contingency table for classification of the validation set samples by SIMCA.

		Predicted dye				
		Yellow 4GN	Bordeaux B	Blue 2R	Red 2B	Without classification
Real dye	Yellow 4GN	6	0	0	0	0
	Bordeaux B	0	6	0	0	0
	Blue 2R	0	0	7	0	0
	Red 2B	0	0	0	5	1

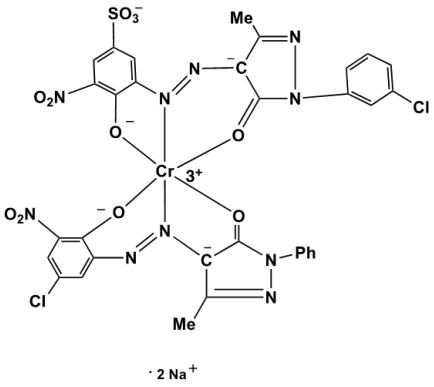
Figures

Figure 1. Chemical structures of the acidic dyes of interest.

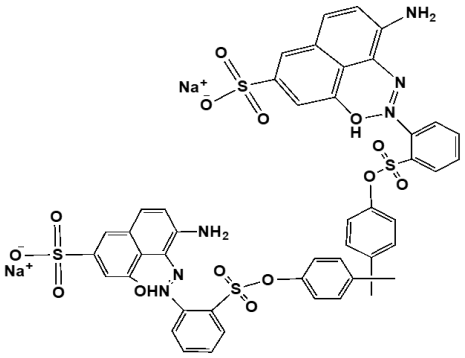
Blue 2R, CAS 70209-96-0



Red 2B, CAS 72017-66-4



Yellow 4GN, CAS 72479-28-8



Bordeaux B, CAS 52333-30-9

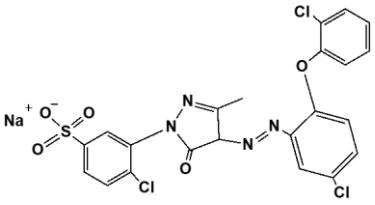


Figure 2. Influence of pH on Yellow 4GN in aqueous solution: a) absorption spectra, and b) chromatic coordinates of the CIELab color space.

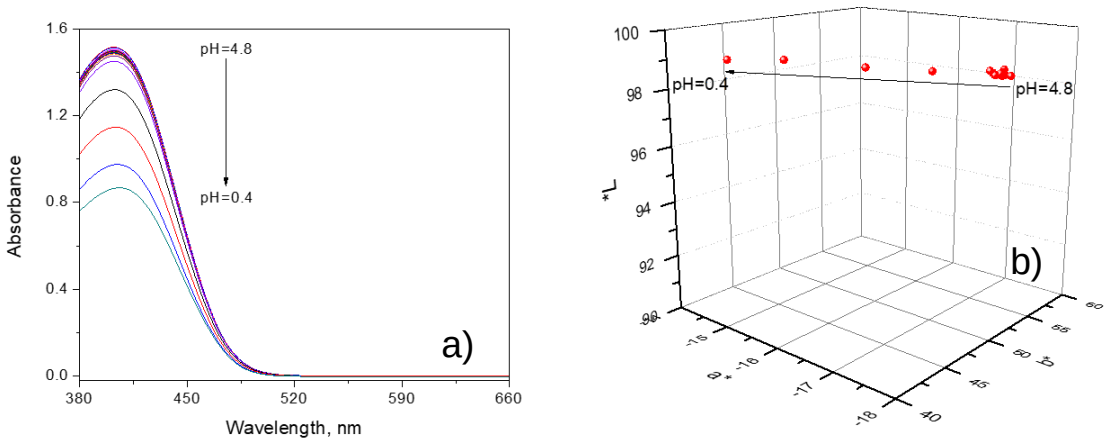


Figure 3. Influence of pH on Blue 2R in aqueous solution: a) absorption spectra, and b) chromatic coordinates of the CIELab color space.

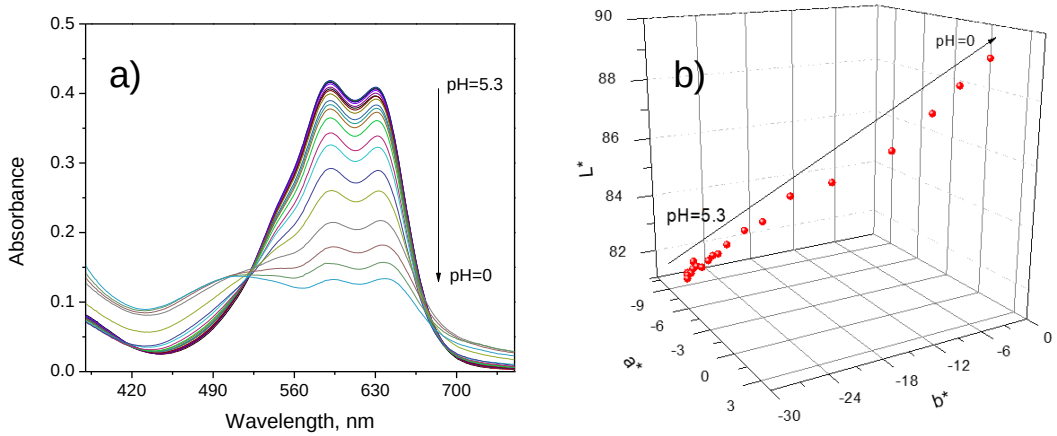


Figure 4. Influence of pH on Red 2B in aqueous solution: a) absorption spectra, and b) chromatic coordinates of the CIELab color space.

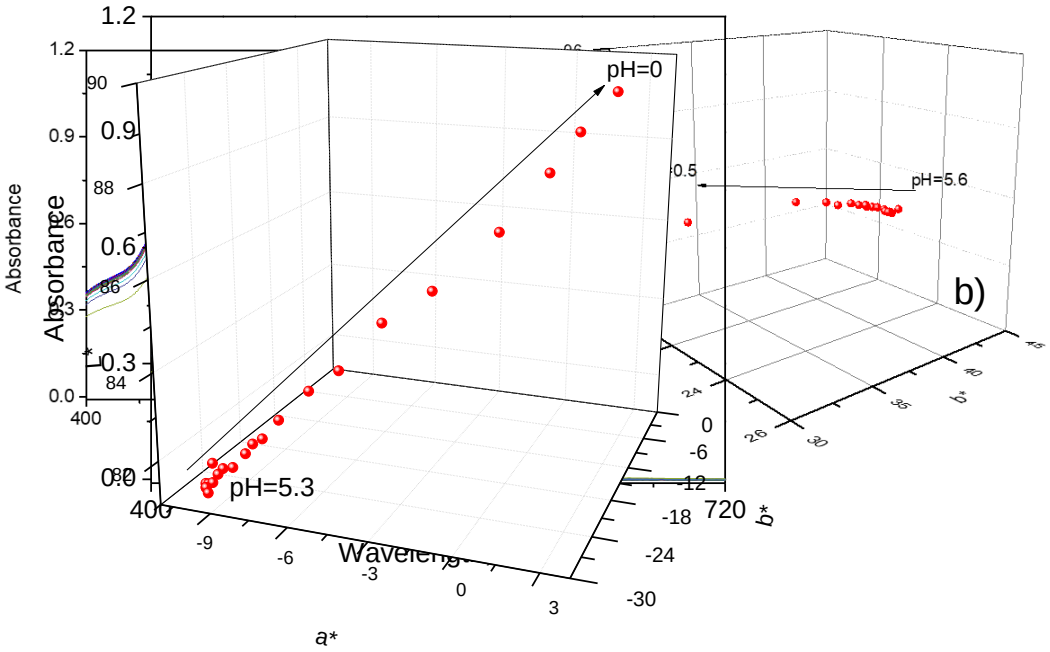


Figure 5. Influence of pH on Bordeaux B in aqueous solution: a) absorption spectra, and b) chromatic coordinates of the CIELab color space.

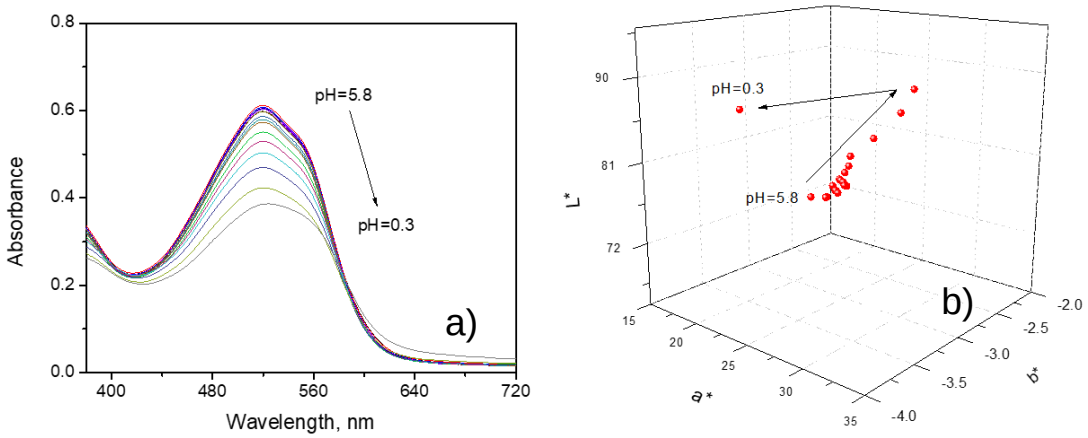


Figure 6. Screening of dyes by SIMCA in a) training and b) validation series.