

## Chemometrics in analytical chemistry – an overview of applications from 2014 to 2018

Mônica Cardoso Santos<sup>1</sup>, Paloma Andrade Martins Nascimento<sup>2</sup>, Wesley Nascimento Guedes<sup>1</sup>, Edénir Rodrigues Pereira-Filho<sup>3</sup>, Érica Regina Filletti<sup>1</sup>, Fabíola Manhas Verbi Pereira<sup>1+</sup>

1 São Paulo State University (UNESP), Institute of Chemistry, 55 Professor Francisco Degni St, Araraquara, São Paulo, Brazil

2 São Paulo State University (UNESP), Faculty of Pharmaceutical Sciences, Km 1 Araraquara-Jaú Hw, Araraquara, São Paulo, Brazil

3 Federal University of São Carlos (UFSCar), Department of Chemistry, Km 235 Washington Luís Hw, São Carlos, São Paulo, Brazil

+Corresponding author: Fabíola Manhas Verbi Pereira, Phone: +55 16 3301 9609, email address: [fabiola.verbi@unesp.br](mailto:fabiola.verbi@unesp.br)

### ARTICLE INFO

#### Article history:

Received: December 2, 2018

Accepted: April 2, 2019

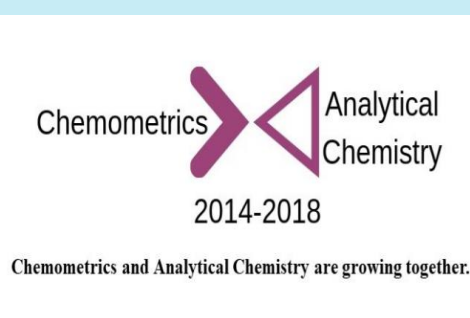
Published: April 25, 2019

#### Keywords:

1. chemometrics

2. analytical chemistry

**ABSTRACT:** A compilation of papers published between 2014 and 2018 was evaluated. Many papers related to multivariate calibration and classification have been reported, as well as, design of experiments applications and artificial intelligence methods. Some applications were highlighted, as medical and pharmaceutical, food analysis, fuels, biological and forensic for the chemometric techniques on this review. Most studies are related to developing methods for practical solutions in industry or routine analysis. A promising scenario is shown considering the number of published papers: a total of 832 for this period using the keywords, multivariate classification, multivariate calibration, analysis, chemometrics, prediction, analytical chemistry, artificial neural networks (ANN), design of experiments (DoE) and factorial design. An useful overview for Analytical Chemistry researchers' combined with Chemometrics is presented in this review.



### CONTENTS

#### 1. Introduction

#### 2. Materials and methods

##### 2.1. Medical and pharmaceutical applications

##### 2.2 Food analysis

##### 2.3 Fuel

#### 3. Results and discussion

##### 3.1 Biological samples

##### 3.2 Food analysis

##### 3.3 Forensic purposes

##### 3.4 Fuels

#### 4. Artificial intelligence methods for multiple questions

#### 5. Design of Experiments (DoE) and response surface methodology (RSM)

#### 6. Conclusions

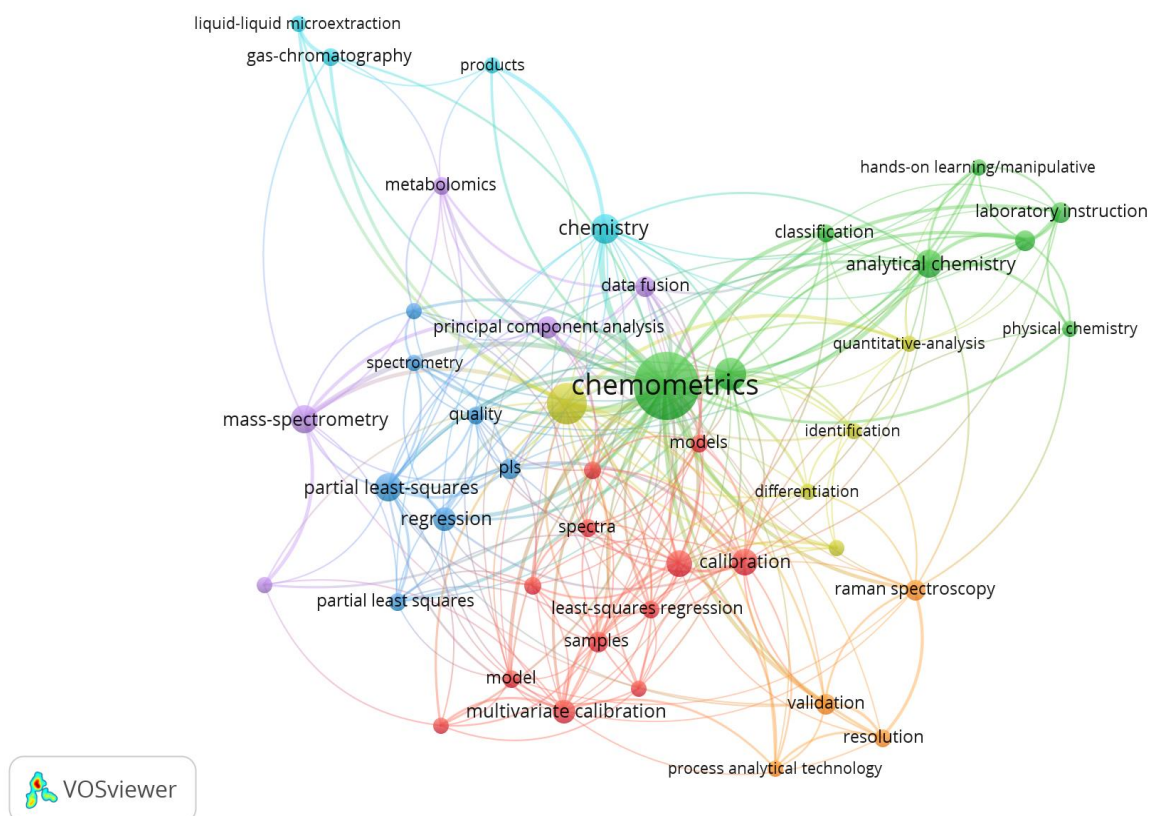
#### 7. Acknowledgments

### 8. References

#### 1. Introduction

With the continuous technological progress of instrumental techniques for analytical purposes, multivariate methods applied to chemical data are mandatory in several applications. Chemometry is a prominent area dedicated to developing multivariate strategies for chemical data evaluation and interpretation<sup>1,2</sup>. Figure 1 shows the main subjects related to Chemometric techniques according to the most repeated words in a bibliographic searching performed in Science Citation Index Expanded (SCI-E) in Clarivate Analytics' ISI – Web of Science<sup>®</sup>, on June 21, 2018.





**Figure 1.** Bibliometric map generated using the more repeated keywords in the authors' search for the period 2014-2018.

The main topics revealed by the search can be verified in the color clusters of Figure 1, as follows: in green symbols classification and educational purposes (“hands-on-learning”, “laboratory instruction” as keywords), light blue circles denote chemistry and separation techniques (chromatography), purple circles represent sophisticated instrumentation as mass-spectrometry, metabolomics, and principal component analysis (PCA), dark blue symbols are related to regression and quality control, as well as the red ones are linked to calibration models and spectral techniques, the process analytical technology (PAT) and identification/differentiation issues appear in the searching as orange and yellow circles, respectively. In the center of the clustering, chemometrics is surrounded by all those words and connected to analytical chemistry.

A total of 832 occurrences were reported by the authors' during the search. Most studies are related to published papers, followed by reviews and proceedings in minor scale. From this amount, studies involving multivariate classification as topic has 390 papers, multivariate calibration 209,

design of experiments (DoE) 136 and artificial neural networks (ANN) 97. It is clear that chemometrics and analytical chemistry are growing together, and the goal of this review was to make a concise compilation of representative studies for the area between 2014 and 2018. Several topics were discussed by using key papers mentioned along the text.

## 2. Materials and methods

### 2.1. Medical and pharmaceutical applications

The use of chemometrics to study multifunctional indole alkaloids from *Psychotria nemorosa* (Palicourea comb. nov.) was developed by Júnior *et al.*<sup>3</sup>. The techniques PCA, partial least squares (PLS) and orthogonal PLS1 (O-PLS1) were helpful for modelling the activities using ultra-high-performance liquid chromatography-diode array detector (UPLC-DAD) data as fingerprint<sup>3</sup>. Ultraviolet spectral data from monoclonal antibodies compounded into hospital pharmacy were evaluated to build models using PLS-discriminant analysis (PLS-DA). The

challenge was to identify the monoclonal antibodies after compounding and just before administration to the patient for quality control at the hospital<sup>4</sup>. The linear sweep voltammograms generated from integration of three commercial screen-printed electrodes into a voltammetry electronic tongue was combined with PLS to predict concentrations of cysteine, glutathione and homocysteine. The authors applied genetic algorithm and interval PLS (i-PLS) for selection of variables<sup>5</sup>. It was possible to monitor the extraction of indole alkaloids from the toad skin<sup>6</sup> and the quantification of geniposide in *Gardenia jasminoides* fruit (Chinese medicinal herb)<sup>7</sup> was performed using near infrared (NIR) spectroscopy and PLS regression (PLSR). The extraction of indole alkaloids from the toad skin showed a coefficient of determination ( $R^2$ ) value of 0.99 and root mean square error of cross validation (RMSECV) of 8.26 mg mL<sup>-1</sup><sup>6</sup> and the correlation values to quantify the concentration of geniposide in *Gardenia* were between 0.92 – 0.99 and 0.32 - 1.66 mg mL<sup>-1</sup> for RMSECV<sup>7</sup>. In order to evaluate the anti-inflammatory properties of extracts from Honeysuckle, Fourier Transform Infrared (FTIR) spectroscopy in combination with PLSR<sup>8</sup> were used.

Junior *et al.*<sup>9</sup> used NIR and performed a comparison between two multivariate calibration techniques, PLSR and multiple linear regression (MLR) to monitor the quality control of tablets, and the PLSR model present better results than MLR. Raman spectral imaging technique was also used to predict active ingredient pharmaceutical concentration in microtablets<sup>10</sup>, in oral suspensions<sup>11</sup>, intravaginal rings<sup>12</sup>, and in quantitative assessment of pharmaceutical tablets<sup>13</sup>. The PLSR and least squares-support vector machine (LS-SVM) models exhibited excellent prediction abilities for active pharmaceutical ingredient in microtablets with a correlation higher than 0.98 and root mean square error of prediction (RMSEP) of 2.35% (w/w) for PLSR and 2.09% (w/w) for LS-SVM<sup>10</sup>. Mazurek and Szostak<sup>11</sup>, using PLSR model, obtained RMSE of calibration (RMSEC) from 1.59 to 2.92% for oral suspensions<sup>11</sup> and Lyndgaard *et al.*<sup>12</sup> achieved RMSECV of 1.82% (w/w) and correlation of 0.99 for intravaginal rings using PLSR model. Then, Sparén *et al.*<sup>13</sup> in quantitative assessment of tablets obtained RMSE of prediction (RMSEP) from 0.54 to 3.04% (w/w). The ultraviolet visible (UV-Vis) spectroscopy and chemometric models also

allowed determination the presence of impurities in tablets<sup>14</sup>. Support vector regression (SVR) model was more accurate when compared to PLSR model with values for RMSEP between 0.18 – 0.27 µg mL<sup>-1</sup> while the values for PLSR were RMSEP from 0.27 to 0.30 µg mL<sup>-1</sup><sup>14</sup>. The synergy PLS (siPLS) model combined with FTIR showed the best results compared to the full spectrum PLS model for the quantification of active ingredient pharmaceutical in tablets which values highlight by authors was RMSECV of 15.28 mg g<sup>-1</sup> and  $R^2$  of 0.99 for sulphamethoxazole and RMSECV of 7.52 mg g<sup>-1</sup> and correlation of 0.98 for trimethoprim<sup>15</sup>. Cantor *et al.*<sup>16</sup> evaluated the ability of different analytical techniques such X ray diffraction, FTIR, Raman and NIR to detect melamine levels in gelatin (pharmaceutical component). The authors highlighted NIR technique and PLSR model that achieved the best results with RMSEP of 2.5% (w/w)<sup>16</sup>, besides to quantify the presence of melamine at levels as low as 1.0% (w/w).

The NIR spectroscopy and PLSR were used to determine protein nitrogen in yellow fever vaccine<sup>17</sup>. Haghghi, Hemmateenejad and Shamsipur<sup>18</sup> using fluorescence spectroscopy and multivariate curve resolution alternative least square (MCR-ALS) were able to predict enantiomeric excess of some amino acids in frozen plasma which RMSEP values lower than 0.07% and R higher than 0.97. PLS was applied in NIR spectra to analyze the indicator of  $\beta$ -thalassemia in human hemolysate samples<sup>19</sup>. FTIR and multivariate calibration method were used to identify three group of clinical bacterial pathogens in human serum<sup>20</sup>. Both models, PCA and PLS-DA, revealing a feasible tool for discrimination of the samples with percent accuracy of 100%<sup>20</sup>.

## 2.2 Food analysis

NIR and PLSR were used to measure both physicochemical<sup>21</sup> and the antioxidant properties of honey<sup>22</sup>, as well to detect and to quantify adulterant in honey samples of stingless bees<sup>23</sup>. The PLSR model presented a RMSECV between 0.008 – 1.47 and correlation values from 0.90 to 0.98 for physicochemical properties<sup>21</sup>, while the antioxidant properties of honeys from different botanical origin showed high correlation varying from of 0.95-0.96 and RMSECV range from 0.92 to 13.60<sup>22</sup>. Se *et al.*<sup>23</sup> obtained values for prediction of adulteration in honey samples of stingless bees

with relative errors up 0.58 to 1.49%. The three-dimensional fluorescence spectroscopy technique combined with multivariate calibration was also used for detecting the presence of adulterants in honey<sup>24</sup>.

A near infrared hyperspectral imaging system was used for predicting viability and vigor in muskmelon seeds using PLS-DA. Variable selection methods, such as variable important projection (VIP), selectivity ratio (SR) and significance multivariate correlation (sMC) were also applied to select the most effective wavelengths<sup>25</sup>.

In order to determine oxidation parameters in frying canola oil, Talpur *et al.*<sup>26</sup> employed FTIR technique and PLSR model, while Uncu and Ozen<sup>27</sup> studied various chemical parameters in olive oils and Wu *et al.* evaluated rapeseed oil adulteration<sup>28</sup>. NIR and PLSR models were also used to study oxidation of edible oils<sup>29</sup>, to detect adulteration in extra virgin olive oils<sup>30</sup>, and to monitor vegetable oils quality<sup>31</sup>. The fluorescence spectroscopy and different multivariate calibration methods, PLSR, PCR and MLR were applied for detection of adulteration of sesame oil. The best results highlighted by the authors were using the PLSR model which figures of merit were RMSECV of 0.65 for 80 nm wavelength intervals<sup>32</sup>.

The applicability of FTIR was used as tool to determine of total sugar content in soy-based drinks<sup>33</sup> and for measuring the concentration of nutrients in grapevine petioles<sup>34</sup>. For the soy-based drinks, two models were applied for different spectral band: (1) iPLS and (2) siPLS, and both presented high coefficient of correlation (R) (0.98).

Milk samples were evaluated using NIR and PLS-DA to detect contamination by *Salmonella*<sup>35</sup>. Mid-infrared (MIR) spectroscopy and PLSR were applied for determining the residues of tetracycline in cow's milk<sup>36</sup> and for detecting and quantifying the adulteration in milk powder<sup>37</sup>. Augusto *et al.* used laser-induced breakdown spectroscopy (LIBS) data and PLSR model for the direct determination of mineral elements, such as Ca, K and Mg in powdered milk and dietary supplements<sup>38</sup>.

### 2.3 Fuel

Data acquired using C-13 nuclear magnetic resonance (NMR) spectroscopy were evaluated using SVR with variable selection by genetic

algorithm (GA) to determine the contents of saturated, aromatic, and polar compounds in crude oil. The use of small amounts of samples was highlighted by the authors<sup>39</sup>. Sulfur content in petroleum derivatives was determined using MIR spectroscopy and PLS in association with variable selection methods, among them iPLS, siPLS, uninformative variable elimination (UVE), and GA<sup>40</sup>. The UV spectroscopy integrated different chemometric strategies for quantifying the hydrocarbon contents in fuel oil samples<sup>41</sup>.

In order to quantify the presence of adulterants in gasoline, Raman spectroscopy was used, and calibration models were constructed using PLSR, iPLS and PLS-GA which correlations were between 0.80 and 0.99<sup>42</sup>. Gas chromatography data were used to build a PLSR model for quality assessment of gasoline, and low errors from 0.005 to 0.010% were obtained<sup>43</sup>. Dadson, Pandam and Asiedu evaluated adulterants in gasoline, such as kerosene, diesel, naphta using a multivariate model with FTIR data and PLSR<sup>44</sup>. Mabood *et al.*<sup>45</sup> also applied PLSR model to detect and estimate gasoline adulteration by NIR spectroscopy.

## 3. Results and discussion

### 3.1 Biological samples

Neves *et al.*<sup>46</sup> used mass spectrometry coupled with PCA, GA with support vector machine (SVM), linear discriminant analysis (LDA) and quadratic discriminant analysis (QDA) as an untargeted lipidomic approach to classify 76 blood plasma samples into negative for intraepithelial lesion or malignancy and squamous intraepithelial lesion. The PCA-SVM models provided the best classification results, achieving values of 80% and 83% for the sensitivity and specificity, respectively. Santos *et al.*<sup>47</sup> proposed a method to determine 11 polycyclic aromatic hydrocarbons (PAHs) in urine samples based on the coupling of a programmed temperature vaporizer (PTV) with a quadrupole mass spectrometer (qMS) instrument, via a deactivated fused silica tubing. The authors used PLS-DA, LDA, soft independent modeling of class analogy (SIMCA) and SVM to classify the samples according to the presence or absence of the PAHs. Gilany *et al.*<sup>48</sup> investigated untargeted metabolomic profiling of the seminal plasma in non-obstructive azoospermia men using gas chromatography-mass spectrometry (GC-MS) and QDA chemometric technique to implement on total

ion chromatograms for identification of discriminatory retention times. The receiver operating characteristic (ROC) curves for these classification models presented 88% of accuracy for the discrimination of 36 metabolites that may be considered discriminatory biomarkers for different groups in non-obstructive azoospermia. Boll *et al.*<sup>49</sup> developed a method using attenuated total reflection (ATR) FTIR spectroscopy combined with PLS-DA model that allowed the discrimination between dyed and non-dyed hair of 380 hair samples. Monakhova, Diehl and Fareed<sup>50</sup> used PCA, factor discriminant analysis (FDA), PLS-DA and LDA combined with high resolution (600 MHz) NMR spectroscopy data to distinguish 102 authentic samples of heparin and low-molecular weight heparins produced from porcine, bovine and ovine mucosal tissues as well as their blends.

### 3.2. Food analysis

Chen *et al.*<sup>51</sup> detected 101 volatile compounds in 70 Chinese vinegars through GC-MS and used chemometric techniques such as clusters analysis, LDA, k nearest neighbor (kNN), GA-ANN to classify these compounds. Giannetti *et al.*<sup>52</sup> evaluated flavor composition of apple cultivars grown in the Northeast Italy through different cultivation forms. The authors used head space-solid phase micro extraction/gas chromatography mass spectrometry (HS-SPME/GC-MS) for the analysis of volatile fraction and PLS-DA model to classify 42 apples varieties. The HS-SPME/GC-MS technique identified 118 volatile compounds and the PLS-DA model classified apples based on their different geographical origin or growing conditions, providing accuracy of 88-100%. Cuevas *et al.*<sup>53</sup> applied data fusion (mid-level) using high-performance liquid chromatography high resolution mass spectrometric (HPLC-HR-MS) and HS-SPME/GC-MS to obtain data sets and PLS-DA model for the discrimination between 19 samples of organic and conventional orange juices. The HPLC-HR-MS and HS-SPME/GC-MS techniques identified some flavonoids, fatty acids, aldehydes and esters as potential markers involved in the discrimination of organic juices. Wen *et al.*<sup>54</sup> characterized the phenolic and proline profiles of four types of unifloral honey collected from beekeepers in China. LDA results showed overall accuracy of 99% and 94% in correct classification of differentiate the floral origin of four Chinese

honeys. Portarena, Baldacchini and Brugnoli<sup>55</sup> proposed a method to discriminate 38 extra-virgin olive oils from seven regions along the Italian coasts from determination of the isotopic composition and the carotenoid contents. The authors used isotope ratio mass spectrometry (IRMS) and resonant Raman spectroscopy (RRS) to determine isotopic composition and the carotenoid contents, respectively. The LDA model presented 50 – 100% of correct classification for the seven regions, achieving 82% of accuracy for validation set.

Márquez *et al.*<sup>56</sup> used NIR and FT-Raman data to perform the data fusion in mid and high-level using SIMCA model for classifying the unadulterated and adulterated hazelnut samples. Rodrigues Júnior *et al.*<sup>57</sup> applied PLS-DA for discrimination of lactose-free samples and classification of adulterated and unadulterated milk powder samples by FT-Raman spectroscopy. Wakholi *et al.*<sup>58</sup> used short-wave IR (SWIR) hyperspectral imaging coupled with LDA, PLS-DA and SVM models for classification of corn seeds viable (control group) and corn seeds non-viable (microwave treated). The authors used a total of 200 samples of each type of corn: white, yellow and purple. Sandasi *et al.*<sup>59</sup> applied a fast and non-destructive method for the quality control of herbal tea blends. The authors used hyperspectral images obtained from SWIR combined with PLS-DA to classify herbal tea blends of the species *Sceletium tortuosum* and *Cyclopia genistoides*.

Kimuli *et al.*<sup>60</sup> also used SWIR hyperspectral imaging system (1000 – 2500 nm) to detect aflatoxin B<sub>1</sub> combined with PCA, PLS-DA and FDA techniques to explore and classify maize kernels of four varieties from different States of the USA. Cortés *et al.*<sup>61</sup> discriminated two nectarine varieties in different spectral ranges of NIR and Vis-NIR using PLS-DA and LDA models. Shrestha *et al.*<sup>62</sup> used NIR hyperspectral imaging data from 975 to 2500 nm and PLS-DA to investigate seed quality parameters such as year of production and variety in tomato seed lots. Santos, Pereira-Filho and Colnago<sup>63</sup> applied <sup>1</sup>H time domain nuclear magnetic resonance (TD-NMR) combined with multivariate analysis for identifying and quantifying the adulteration in 78 samples of milk using whey, urea, hydrogen peroxide and synthetic material (urine or milk).

### 3.3 Forensic purposes

Chen *et al.*<sup>64</sup> used NIR combined with hierarchical cluster analysis (HCA), PLS-DA and SVM to authenticate stamps of 12 seals on a Chinese traditional painting. The results obtained were 93% and 100% of accuracy for the PLS-DA and SVM model in validation set, respectively. Oliveira *et al.*<sup>65</sup> discriminated banknotes authenticity of Brazilian Real (R\$ 20, R\$ 50 and R\$ 100) using a portable NIR spectrometer combined with SIMCA and successive projections algorithm-LDA (SPA-LDA) models. Martins *et al.*<sup>66</sup> proposed a method for direct analysis of seven whisky brands. The authors used UV-Vis spectroscopy combined with PLS-DA to classify 164 genuine and 73 false samples. The PLS-DA corrected classified 99% and 93% for genuine and false samples, respectively. Costa *et al.*<sup>67</sup> identified six polymers using different ratios of emission lines and molecular bands (C/H, C/C<sub>2</sub>, C/N, C/O, H/C<sub>2</sub>, H/N, H/O and O/N) obtained with LIBS. SIMCA and kNN models were computed using 477 samples divided between calibration and validation sets. Both models achieved accuracy of 98% (kNN) and 92% (SIMCA).

### 3.4 Fuels

Silva *et al.*<sup>68</sup> developed a method to classify gasoline according to its origin (Brazil, Venezuela and Peru), using IR spectroscopy and multivariate classification. A set of 126 gasoline samples: 56 Brazilian, 66 Venezuelan, and 4 Peruvian, was analyzed. The spectra were standardized using the direct standardization method achieving 100% of correct predictions. In another study, Silva *et al.*<sup>69</sup> proposed the use of IR spectroscopy and LDA and PLS-DA for classification of gasoline with or without dispersant and detergent additives in 125 samples. The authors selected the variables for the multivariate models using stepwise (SW), successive projections algorithm (SPA) and GA algorithms. Sinkov, Sandercock and Harynuk<sup>70</sup> developed PLS-DA and SIMCA models for classifying the levels of gasoline in casework arson samples based on GC-MS data.

## 4. Artificial intelligence methods for multiple questions

An interesting approach for ANN is to handle the samples more efficiently to achieve streamlined processes. Kovalishyn and Poda<sup>71</sup> created a new variable selection method for ANN known as batch pruning algorithm (BPA), being faster and more efficient than traditional methods. ANN was used for Sobol sensitivity estimation<sup>72</sup> evidencing as advantage, the reduction of computational costs. Still, Bian *et al.*<sup>73</sup> allied the advantages of linear and non-linear methods in a novel algorithm called the extreme learning machine (ELM) and compared with PCR, PLS, SVR and back propagation algorithm-ANN (BP-ANN) by three NIR spectral datasets: diesel fuel, ternary mixture and blood, showing that ELM presented the best performance in the spectral quantitative analysis of complex samples.

NIR spectra data acquired for beers from three types of fermentation was used to evaluate foam and color-related parameters<sup>74</sup>, using a robotic pourer and chemical fingerprinting. Results from NIR were used to create PLS and ANN models to predict four parameters such as pH, alcohol, Brix and maximum volume of foam. The ANN was implemented using the Levenberg–Marquardt training algorithm, being able to create more accurate models than PLS<sup>74</sup>. Oliveira *et al.*<sup>75</sup> determined protein concentration, foam stability, haze, color, total acidity (TA), alcohol content, and bitterness in Ale beers properties using UV-Vis spectra in combination with PCR and also with ANN models.

Many authors also compared ANN and other chemometric methods, such as studies of Das *et al.*<sup>76</sup> that modeled variations in sucrose, reducing and total sugar content due to water-deficit stress in rice leaves using Vis, NIR and SWIR spectroscopies. They tested the following multivariate techniques, ANN feed-forward model, multivariate adaptive regression splines (MARS), MLR, PLSR, random forest regression and SVM. The variables that affect the performance of ANNs and PLS for spectral interference correction are random noise level, intensity ratio, peak separation and wavelength shift. The results showed that ANNs and PLS are about equally as effective for spectral interference correction<sup>77,78</sup>. A portable artificial olfaction system for real-time monitoring of black tea fermentation was developed, based on the combination of the kNN and adaptive boosting, namely kNN-AdaBoost with discrimination rate of 100% of correct predictions<sup>79</sup>. The performance of decision tree, kNN, naive bayesian, SVM and ANN

for to classification of orange growing locations based on the NIR using data mining was evaluated<sup>80</sup>. The experimental results showed that the juice NIR spectra is the most suitable dataset for identifying the orange growing locations, and the decision tree is the best and most stable classifier<sup>80</sup>, with the average prediction rate of 97%.

In relation to herbal products, Wang *et al.*<sup>81</sup> used LIBS combined with PCA and BP-ANN to classify Chinese herbal medicine with 99.9% classification accuracy of three types of herbal products, roots of *Angelica pubescens*, *Codonopsis pilosula*, and *Ligusticum wallichii*. The authors established<sup>82</sup> quality control markers for Chinese herbal medicines using BP-ANN; and the study of Ding *et al.*<sup>83</sup> developed a method to improve the markers to Q-markers in Chinese herbal medicines quality management, using PLS-DA for screening analysis of the chemical markers and identification of herbal origin. The BP-ANN algorithm was used to clarify the non-linear relationship between the Q-markers and their integral anti-inflammation effect. Still in relation to herbal medicines, Ito *et al.*<sup>84</sup> developed a model based on an ANN to quantify proteolytic and amylolytic enzymes using UV-Vis spectra of diluted samples from a particular solid-state fermentation in wheat bran, soybean meal, type II wheat flour and sugarcane bagasse.

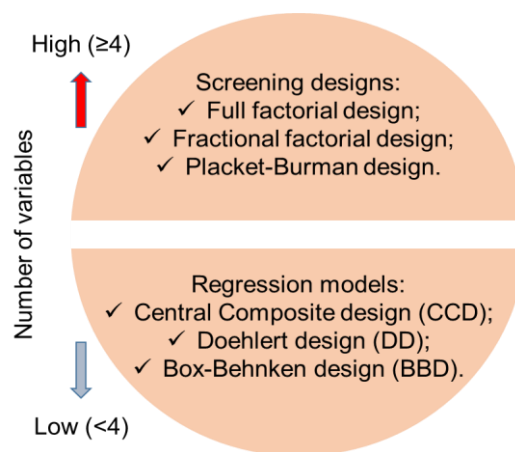
Li *et al.*<sup>85</sup> combined GA with ANN to determine the elements copper and vanadium in steel samples with satisfactory quantitative results. Zhang *et al.*<sup>86</sup> developed a method combining GA, PCA and ANN to select spectral segments from the original spectra to improve the LIBS performance and proved that use only a fixed-length segment appropriate provides better results than selecting the entire spectral range.

Gurbanov, Bilgin and Severcan<sup>87</sup> used ANN to evaluate the secondary structural variations in the diabetic kidney cell membrane proteins based in ATR-FTIR spectroscopy analyzing the effects of selenium on diabetes. Hasanjani and Sohrabi<sup>88</sup> used UV-Vis spectroscopic and back-propagation algorithm ANN to predict fluoxetine and sertraline in tablets. Guo *et al.*<sup>89</sup> developed a kinetic spectrofluorometric method for the analysis of sibutramine, indapamide, and hydrochlorothiazide, very common in weight-reducing health foods, and the data of the mixtures were processed by parallel factor analysis (PARAFAC), PLS, PCR and radial basis function-artificial neural network (RBF-ANN).

## 5. Design of Experiments (DoE) and response surface methodology (RSM)

Before the application of any analytical method, it is necessary to optimize some instrumental conditions, named variables. These variables can be related to sample preparation procedure, as volumes of reagents, pressure and temperature or instrumental parameters, as wavelength, power, among others. The goal in several cases is always related to obtain a condition with high analytical signal, signal-to-noise ratio (SNR), and signal-to-background (SBR). In addition, analysts can be also interested in analytical methods with low relative standard deviation, limits of detection, and quantification, and reduced cost per determination. In order to achieve these goals DoE can be applied in almost any type of problem<sup>90-92</sup>.

DoE can be defined as a group of tools that applies statistical and mathematical knowledge to optimize a system. Several strategies can be used as central composite design (CCD)<sup>93</sup>, Doehlert (DD)<sup>94</sup> and, Box-Behnken designs (BBD)<sup>95</sup>. If the goal is only to identify the most important variables, full factorial<sup>96</sup> or fractional factorial<sup>97</sup> or Plackett-Burman<sup>98</sup> designs can be also applied. Figure 2 shows the several possibilities related to the application of DoE in any type of problem.



**Figure 2.** DoE characteristics for optimization.

The applicability and advantages of DoE in analytical chemistry are out of context and hundreds of studies can be easily searched in the scientific literature. In the time span selected for this review we found 125 papers using only the combination of the words “DoE”, “Factorial Design”, “Chemistry” and “Analytical chemistry”.

The scientific literature presents several reviews and tutorial reviews about the use of DoE<sup>90, 92</sup>. In this case, the fundamentals about the use of DoE will be not discussed in this review, and only

applications will be described and discussed. Table 1 shows the main applications observed and several remarks.

**Table 1.** Selected analytical problems that applied DoE.

Analytical chemistry type of application	Problems optimized	References
Chromatography	<ul style="list-style-type: none"> <li>- Matrix effects correction in liquid chromatography</li> <li>- Vitamin D3 determination</li> <li>- Synthesis of hydrogel</li> <li>- Single use bioreactors</li> <li>- Supercritical fluid</li> </ul>	99-103
Environmental	<ul style="list-style-type: none"> <li>- Determination of organic pollutants (<math>\beta</math>-cyclodextrin)</li> <li>- Occupational exposition analysis</li> <li>- NO<sub>2</sub> detection using Raman spectroscopy</li> <li>- N<sub>2</sub>O and SF<sub>6</sub> determination</li> <li>- Cloud point extraction</li> <li>- N and S deposition</li> <li>- Cu detection in river water</li> </ul>	104-106
Material	<ul style="list-style-type: none"> <li>- Formation of Ca neptunatos</li> <li>- Signal processing in Raman</li> <li>- Steel production</li> <li>- Environmental aging simulation</li> <li>- Materials proposition for environmental analysis</li> </ul>	107,108
Bioanalytical	<ul style="list-style-type: none"> <li>- Protein translation</li> <li>- Seed growth</li> </ul>	109
Pharmaceutical	<ul style="list-style-type: none"> <li>- Biosensor development</li> <li>- Quantification of cefazolin</li> <li>- Drugs stability</li> <li>- Acetylcholinesterase determination</li> </ul>	110
Electroanalytical	<ul style="list-style-type: none"> <li>- Instrumental conditions of anodic stripping voltammetry for Cu<sup>2+</sup> determination in water</li> </ul>	111
Food	<ul style="list-style-type: none"> <li>- Ascorbic acid determination</li> <li>- Food safety</li> </ul>	112
NMR	<ul style="list-style-type: none"> <li>- Metabolomics</li> <li>- <i>In vitro</i> mutagenicity</li> </ul>	113,114

## 6. Conclusions

According to authors' search of this review, the most developed methods combined with chemometrics is for food analysis, undoubtedly. In general, the goal of the studies was to improve speediness in the analysis and, to reduce the number of steps in the analytical method. Fuel samples were also a highlighted topic in analytical

chemistry. Sophisticated techniques considering mass spectrometry and chromatography allied to chemometrics have been increasing their applications in the area. Classification techniques have the majority of the papers, following by multivariate calibration, DoE and ANN. Chemometrics and analytical chemistry is a powerful combination to improve robustness, analytical frequency and practicality for methods.



## 7. Acknowledgements

The authors are grateful to the São Paulo Research Foundation (FAPESP grants number 2018/18212-8 and 2019/01102-8) and the National Council for Scientific and Technological Development (CNPq). This study was financed in part by the Coordination for the Improvement of Higher Education Personnel (CAPES) – Finance Code 001.

## 8. References

- [1] Olivieri, A. C., Introduction to multivariate calibration: A practical approach, Springer International Publishing, Switzerland, 1st ed., 2018, XVII-243. <https://doi.org/10.1007/978-3-319-97097-4>.
- [2] Lavine, B. K., Rayens, W. S., Classification: basic concepts, in: Brown, S. D., Tauler, R., Walczak, B. (Eds.), Comprehensive Chemometrics: Chemical and Biochemical Data Analysis, 3 Elsevier, Amsterdam, The Netherlands, (2009) 507-515.
- [3] Júnior, L. C. K., Viaene, J., Tuenter, E., Salton, J., Gasper, A. L., Apers, S., Andries, J. P. M., Pieters, L., Henriques, A. T., Heyden, Y. V., The use of chemometrics to study multifunctional indole alkaloids from *Psychotria nemorosa* (*Palicourea comb. nov.*). Part II: Indication of peaks related to the inhibition of butyrylcholinesterase and monoamine oxidase-A, *J. Chromatogr. A* 1463 (2016) 71-80. <https://doi.org/10.1016/j.chroma.2016.08.005>.
- [4] Jaccoulet, E., Boccard, J., Taverna, M., Azevedos, A. S., Rudaz, S., Smadja, C., High-throughput identification of monoclonal antibodies after compounding by UV spectroscopy coupled to chemometrics analysis, *Anal. Bioanal. Chem.* 408 (21) (2016) 5915-5924. <https://doi.org/10.1007/s00216-016-9708-4>.
- [5] Pedroza, R. H. P., Serrano, N., Diaz-Cruz, J. M., Arino, C., Esteban, M., Integration of Commercial screen-printed electrodes into a voltammetric electronic tongue for the analysis of aminothiols, *Electroanalysis* 28 (7) (2016) 1570-1577. <https://doi.org/10.1002/elan.201501112>.
- [6] Tao, L., Liu, B., Jin, Y., Sun, D., Liu, X., Chen, Y., Wu, Y., Characterization of Toad Skin for Traditional Chinese Medicine by Near-Infrared Spectroscopy and Chemometrics, *Anal. Lett.* 50 (8) (2017) 1292-1306. <https://doi.org/10.1080/00032719.2016.1220562>.
- [7] Li, J., Xu, B., Zhang, Y., Dai, S., Sun, F., Shi, X., Qiao, Y., Determination of Geniposide in *Gardenia jasminoides* Ellis fruit by near-infrared spectroscopy and chemometrics, *Anal. Lett.* 49 (13) (2016) 2063-2076. <https://doi.org/10.1080/00032719.2015.1130714>.
- [8] Langerodi, R. N., Ortmann, S., Wenzig, E. M. P., Bochkov, V., Zhao, Y. M., Miao, J. H., Saukel, J., Ladurner, A., Heiss, E. H., Dirsch, V. M., Bauer, R., Atanasov A. G., Assessment of anti-inflammatory properties of extracts from Honeysuckle (*Lonicera* sp. L., *Caprifoliaceae*) by ATR-FTIR spectroscopy, *Talanta* 175 (2017) 264-272. <https://doi.org/10.1016/j.talanta.2017.07.045>.
- [9] Junior, S. G., França, L. M., Pimentel, M. F., Albuquerque, M. M., Santana, D. P., Santana, A. K. M., Souza, J. A. L., Simões, S. S., A process analytical technology approach for the production of fixed-dose combination tablets of zidovudine and lamivudine using near infrared spectroscopy and chemical images, *Microchem. J.* 118 (2015) 252-258. <https://doi.org/10.1016/j.microc.2014.07.009>.
- [10] Kandpal, L. M., Cho, B. K., Tewari, J., Gopinathan, N., Raman spectral imaging technique for API detection in pharmaceutical microtablets, *Sens. Actuators B: Chem.* 260 (2018) 213-222. <https://doi.org/10.1016/j.snb.2017.12.178>.
- [11] Mazurek, S., Szostak, R., Quantification of active ingredients in pharmaceutical suspensions by FT Raman spectroscopy, *Vib. Spectrosc.* 93 (2017) 57-64. <https://doi.org/10.1016/j.vibspec.2017.10.003>.
- [12] Lyndgaard, L. B., Spåberg, R., Gilmour, C., Lyndgaard, C. B., Van Den Berg, F., A process analytical approach for quality control of dapivirine in HIV preventive vaginal rings by Raman spectroscopy, *J. Raman Spectrosc.* 45 (2) (2014) 149-156. <https://doi.org/10.1002/jrs.4433>.
- [13] Sparén, A., Hartman, M., Fransson, M., Johansson, J., Svensson, O., Matrix effects in quantitative assessment of pharmaceutical tablets using transmission Raman and near-infrared (NIR) spectroscopy, *Appl. Spectrosc.* 69 (5) (2015) 580-589. <https://doi.org/10.1366/14-07645>.
- [14] Naguib, I. A., Abdelaleem, E. A., Draz, M. E., Zaazaa, H. E., Linear support vector regression and partial least squares chemometric models for determination of Hydrochlorothiazide and Benazepril hydrochloride in presence of related impurities: A comparative study, *Spectrochim. Acta Part A: Molecular and Biomolecular Spectroscopy* 130 (2014) 350-356. <https://doi.org/10.1016/j.saa.2014.04.024>.

- [15] da Silva, F. E. B., Flores, E. M. M., Parisotto, G., Müller, E. I., Ferrão, M. F., Green method by diffuse reflectance infrared spectroscopy and spectral region selection for the quantification of sulphamethoxazole and trimethoprim in pharmaceutical formulations, *An. Acad. Bras. Ciênc.* 88 (2016) 1-15. <https://doi.org/10.1590/0001-3765201620150057>.
- [16] Cantor, S. L., Gupta, A., Khan, M. A., Analytical methods for the evaluation of melamine contamination, *J. Pharm. Sci.* 103 (2) (2014) 539-544. <https://doi.org/10.1002/jps.23812>.
- [17] Dabkiewicz, V. E., Abrantes, S. M. P., Cassella, R. J., Development of a non-destructive method for determining protein nitrogen in a yellow fever vaccine by near infrared spectroscopy and multivariate calibration, *Spectrochim. Acta Part A: Molecular and Biomolecular Spectroscopy* 201 (2018) 170-177. <https://doi.org/10.1016/j.saa.2018.04.042>.
- [18] Haghighi, A. N., Hemmateenejad, B., Shamsipur, M., Determination of enantiomeric excess of some amino acids by second-order calibration of kinetic-fluorescence data, *Anal. Biochem.* 550 (2018) 15-26. <https://doi.org/10.1016/j.ab.2018.04.004>.
- [19] Chen, J., Peng, L., Han, Y., Yao, L., Zhang, J., Pan, T., A rapid quantification method for the screening indicator for  $\beta$ -thalassemia with near-infrared spectroscopy, *Spectrochim. Acta Part A: Molecular and Biomolecular Spectroscopy* 193 (2018) 499-506. <https://doi.org/10.1016/j.saa.2017.12.068>.
- [20] Lin, W., Chai, Q., Chen, X., Li, Y., Effective identification of clinical bacterial pathogens by Fourier transform near-infrared spectroscopy, In: *Proc - 2017 Chinese Autom. Congr. CAC* (2017) 4587-4592. <https://doi.org/10.1109/CAC.2017.8243589>.
- [21] Anguebes, F., Pat, L., Ali, B., Guerrero, A., Córdova, A. V., Abatal, M., Garduza, J. P., Application of multivariable analysis and FTIR-ATR spectroscopy to the prediction of properties in campeche honey, *J. Anal. Methods Chem.* 2016 (2016) 1-14. <https://doi.org/10.1155/2016/5427526>.
- [22] Tahir, H. E., Xiaobo, Z., Tinting, S., Jiyong, S., Mariod, A. A., Near-Infrared (NIR) spectroscopy for rapid measurement of antioxidant properties and discrimination of sudanese honeys from different botanical origin, *Food Anal. Methods* 9 (9) (2016) 2631-2641. <https://doi.org/10.1007/s12161-016-0453-2>.
- [23] Se, K. W., Ghoshal, S. K., Wahab, R. A., Ibrahim, R. K. R., Lani, M. N., A simple approach for rapid detection and quantification of adulterants in stingless bees (*Heterotrigona itama*) honey, *Food Res. Int.* 105 (2018) 453-460. <https://doi.org/10.1016/j.foodres.2017.11.012>.
- [24] Chen, Q., Qi, S., Li, H., Han, X., Ouyang, Q., Zhao, J., Determination of rice syrup adulterant concentration in honey using three-dimensional fluorescence spectra and multivariate calibrations, *Spectrochim. Acta Part A: Molecular and Biomolecular Spectroscopy* 131 (2014) 177-182. <https://doi.org/10.1016/j.saa.2014.04.071>.
- [25] Kandpal, L. M., Lohumi, S., Kim, M. S., Kang, J. S., Cho, B. K., Near-infrared hyperspectral imaging system coupled with multivariate methods to predict viability and vigor in muskmelon seeds, *Sens. Actuators B: Chem.* 229 (2016) 534-544. <https://doi.org/10.1016/j.snb.2016.02.015>.
- [26] Talpur, M. Y., Hassan, S. S., Sherazi, S. T. H., Mahesar, S. A., Kara, H., Kandhro, A. A., Sirajuddin., A simplified FTIR chemometric method for simultaneous determination of four oxidation parameters of frying canola oil, *Spectrochim. Acta - Part A* 149 (2015) 656-661. <https://doi.org/10.1016/j.saa.2015.04.098>.
- [27] Uncu, O., Ozen, B., Prediction of various chemical parameters of olive oils with Fourier transform infrared spectroscopy, *LWT - Food Sci. Technol.* 63 (2) (2015) 978-984. <https://doi.org/10.1016/j.lwt.2015.05.002>.
- [28] Wu, Z., Li, H., Tu, D., Application of Fourier transform infrared (FT-IR) spectroscopy combined with chemometrics for analysis of rapeseed oil adulterated with refining and purifying waste cooking oil, *Food Anal. Methods* 8 (10) (2015) 2581-2587. <https://doi.org/10.1007/s12161-015-0149-z>.
- [29] Wójcicki, K., Khmelinskii, I., Sikorski, M., Sikorska, E. Near and mid infrared spectroscopy and multivariate data analysis in studies of oxidation of edible oils, *Food Chem.* 187 (2015) 416-423. <https://doi.org/10.1016/j.foodchem.2015.04.046>.
- [30] De Luca, M., Restuccia, D., Lisa, M., Puoci, F., Ragno, G., Chemometric analysis for discrimination of extra virgin olive oils from whole and stoned olive pastes, *Food Chem.* 202 (2016) 432-437. <https://doi.org/10.1016/j.foodchem.2016.02.018>.
- [31] Moreira, S. A., Sarraguça, J., Saraiva, D. F., Carvalho, R., Lopes, J. A., Optimization of NIR spectroscopy based PLSR models for critical properties of vegetable oils used in biodiesel production, *Fuel* 150 (2015) 697-704. <https://doi.org/10.1016/j.fuel.2015.02.082>.
- [32] Temiz, H. T., Tamer, U., Berkkan, A., Boyaci, I. H., Synchronous fluorescence spectroscopy for

- determination of tahini adulteration, *Talanta* 167 (2017) 557-562. <https://doi.org/10.1016/j.talanta.2017.02.044>.
- [33] Rech, A. M., Weiler, F. H., Ferrão, M. F., Determination of total sugar content in soy-based drinks using infrared spectroscopy and chemometrics, *Food Anal. Methods* 11 (7) (2018) 1986-1993. <https://doi.org/10.1007/s12161-018-1170-9>.
- [34] Smith, J. P., Schmidtke, L. M., Müller, M. C., Holzappel, B. P., Measurement of the concentration of nutrients in grapevine petioles by attenuated total reflectance Fourier transform infrared spectroscopy and chemometrics, *Aust. J. Grape Wine Res.* 20 (2) (2014) 299-309. <https://doi.org/10.1111/ajgw.12072>.
- [35] Pereira, J. M., Leme, L. M., Perdoncini, M. R. F. G., Valderrama, P., Março, P. H., Fast Discrimination of milk contaminated with *Salmonella* sp. via near-infrared spectroscopy, *Food Anal. Methods* 11 (7) (2018) 1878-1885. <https://doi.org/10.1007/s12161-017-1090-0>.
- [36] Casarrubias-Torres, L. M., Meza-Márquez, O. G., Osorio-Revilla, G., Gallardo-Velazquez, T., Mid-infrared spectroscopy and multivariate analysis for determination of tetracycline residues in cow's milk, *Acta Vet. Brno.* 87 (2) (2018) 181-188. <https://doi.org/10.2754/avb201887020181>.
- [37] Carvalho, B. M. A., Carvalho, L. M., Coimbra, J. S. R., Minim, L. A., Barcellos, E. S., Júnior, W. F. S., Detmann, E., Carvalho, G. G. P., Rapid detection of whey in milk powder samples by spectrophotometric and multivariate calibration, *Food Chem.* 174 (2015) 1-7. <https://doi.org/10.1016/j.foodchem.2014.11.003>.
- [38] Santos, A. A., Barsanelli, P. L., Pereira, F. M. V., Pereira-Filho, E. R., Calibration strategies for the direct determination of Ca, K, and Mg in commercial samples of powdered milk and solid dietary supplements using laser-induced breakdown spectroscopy (LIBS), *Food Res Int.* 94 (2017) 72-78. <https://doi.org/10.1016/j.foodres.2017.01.027>.
- [39] Filgueiras, P. R., Portela, N. A., Silva, S. R. C., Castro, E. V. R., Oliveira, L. M. S. L., Dias, J. C. M., Neto, A. C., Romão, W., Poppi, R. J., Determination of saturates, aromatics, and polars in crude oil by  $c$ -13 nmr and support vector regression with variable selection by genetic algorithm, *Energy and Fuels* 30 (3) (2016) 1972-1978. <https://doi.org/10.1021/acs.energyfuels.5b02377>.
- [40] Rocha, J. T. C., Oliveira, L. M. S. L., Dias, J. C. M., Pinto, U. B., Marques, M. D. S. P., Oliveira, B. P., Filgueiras, P. R., Castro, E. V. R., Oliveira, M. A. L., Sulfur determination in brazilian petroleum fractions by mid-infrared and near-infrared spectroscopy and partial least squares associated with variable selection methods, *Energy and Fuels* 30 (1) (2016) 698-705. <https://doi.org/10.1021/acs.energyfuels.5b02463>.
- [41] Bian, X., Li, S., Lin, L., Tan, X., Fan, Q., Li, M., High and low frequency unfolded partial least squares regression based on empirical mode decomposition for quantitative analysis of fuel oil samples, *Anal. Chim. Acta* 925 (2016) 16-22. <https://doi.org/10.1016/j.aca.2016.04.029>.
- [42] Ardila, J. A., Luis, F., Soares, F., Antônio, M., Farias, S., Carneiro, R. L., Characterization of gasoline by Raman spectroscopy with chemometric analysis, *Anal. Lett.* 50 (7) (2017) 1126-1138. <https://doi.org/10.1080/00032719.2016.1210616>.
- [43] Parastar, H., Mostafapour, S., Azimi, G., Quality assessment of gasoline using comprehensive two-dimensional gas chromatography combined with unfolded partial least squares: A reliable approach for the detection of gasoline adulteration, *J. Sep. Sci.* 39 (2) (2016) 367-374. <https://doi.org/10.1002/jssc.201500720>.
- [44] Dadson, J., Pandam, S., Asiedu, N., Modeling the characteristics and quantification of adulterants in gasoline using FTIR spectroscopy and chemometric calibrations, *Cogent Chem.* 4 (1) (2018) 1-22. <https://doi.org/10.1080/23312009.2018.1482637>.
- [45] Mabood, F., Gilani, S. A., Albroumi, M., Alameru, S., Nabhani, M. M. O. A., Jabeen, F., Hussain, J., Al Harrasi, A., Boqué, R., Farooq, S., Hamaed, A. M., Naureen, Z., Khan, A., Hussain, Z., Detection and estimation of Super premium 95 gasoline adulteration with Premium 91 gasoline using new NIR spectroscopy combined with multivariate methods, *Fuel* 197 (2017) 388-396. <https://doi.org/10.1016/j.fuel.2017.02.041>.
- [46] Neves, A. C. O., Morais, C. L. M., Mendes, T. P. P., Vaz, B. G., Lima, K. M. G., Mass spectrometry and multivariate analysis to classify cervical intraepithelial neoplasia from blood plasma: an untargeted lipidomic study, *Sci. Rep.* 8 (2018) 3954-3963. <https://doi.org/10.1038/s41598-018-22317-6>.
- [47] Santos, P. M., Sánchez, M. N., Pavón, J. L. P., Cordero, B. M., Quantitative and qualitative analysis of polycyclic aromatic hydrocarbons in urine samples using a non-separative method based on mass spectrometry, *Talanta* 181 (2018) 373-379. <https://doi.org/10.1016/j.talanta.2018.01.032>.
- [48] Gilany, K., Mirzajani, F., Rezadoost, H., Sadeghi, M. R., Amini, M., Untargeted metabolomic profiling of seminal plasma in nonobstructive azoospermia men: A noninvasive detection of spermatogenesis, *Biomed.*

- Chromatogr. 31 (2017) 1-10.  
<https://doi.org/10.1002/bmc.3931>.
- [49] Boll, M. S., Doty, K. C., Wickenheiser, R., Lednev, I. K., Differentiation of hair using ATR FT-IR spectroscopy: a statistical classification of dyed and non-dyed hairs, *Forensic Chem.* 6 (2017) 1-9.  
<https://doi.org/10.1016/j.forc.2017.08.001>.
- [50] Monakhova, Y. B., Diehl, B. W. K., Fareed, J., Authentication of animal origin of heparin and low molecular weight heparin including ovine, porcine and bovine species using 1D NMR spectroscopy and chemometric tools, *J. Pharm. Biomed. Anal.* 149 (2018) 114-119. <https://doi.org/10.1016/j.jpba.2017.10.020>.
- [51] Chen, Y., Bai, Y., Xu, N., Zhou, M., Li, D., Wang, C., Hu, Y., Classification of Chinese vinegars using optimized artificial neural networks by genetic algorithm and other discriminant techniques, *Food Anal. Methods* 10 (2017) 2646-2656.  
<https://doi.org/10.1007/s12161-017-0829-y>.
- [52] Giannetti, V., Mariani, M. B., M., Mannino, P., Marini, F., Volatile fraction analysis by HS-SPME/GC-MS and chemometric modeling for traceability of apples cultivated in the Northeast Italy, *Food Control* 78 (2017) 215-221.  
<https://doi.org/10.1016/j.foodcont.2017.02.036>.
- [53] Cuevas, F. J., Pereira-Caro, G., Moreno-Rojas, J. M., Muñoz-Redondo, J. M., Ruiz-Moreno, M. J., Assessment of premium organic orange juices authenticity using HPLC-HR-MS and HS-SPME-GC-MS combining data fusion and chemometrics, *Food Control* 82 (2017) 203-211.  
<https://doi.org/10.1016/j.foodcont.2017.06.031>.
- [54] Wen, Y.Q., Zhang, J., Li, Y., Chen, L., Zhao, W., Zhou, J., Jin, Y., Characterization of chinese unifloral honeys based on proline and phenolic content as markers of botanical origin, using multivariate analysis, *Molecules* 22 (2017) 735-746.  
<https://doi.org/10.3390/molecules22050735>.
- [55] Portarena, S., Baldacchini, C., Brugnoli, E., Geographical discrimination of extra-virgin olive oils from the Italian coasts by combining stable isotope data and carotenoid content within a multivariate analysis, *Food Chem.* 215 (2017) 1-6.  
<https://doi.org/10.1016/j.foodchem.2016.07.135>.
- [56] Márquez, C., López, M. I., Ruisánchez, I., Callao, M. P., FT-Raman and NIR spectroscopy data fusion strategy for multivariate qualitative analysis of food fraud, *Talanta* 161 (2016) 80-86.  
<https://doi.org/10.1016/j.talanta.2016.08.003>.
- [57] Rodrigues Júnior, P. H., Oliveira, K. S., Almeida, C. E. R., De Oliveira, L. F. C., Stephani, R., Pinto, M. S., Carvalho, A. F., Perrone, Í. T., FT-Raman and chemometric tools for rapid determination of quality parameters in milk powder: Classification of samples for the presence of lactose and fraud detection by addition of maltodextrin, *Food Chem.* 196 (2016) 584-588.  
<https://doi.org/10.1016/j.foodchem.2015.09.055>.
- [58] Wakholi, C., Kandpal, L. M., Lee, H., Bae, H., Park, E., Kim, M. S., Mo, C., Lee, W.-H., Cho, B.-K., Rapid assessment of corn seed viability using short wave infrared line-scan hyperspectral imaging and chemometrics, *Sens. Actuators B. Chem.* 255 (2018) 498-507. <https://doi.org/10.1016/j.snb.2017.08.036>.
- [59] Sandasi, M., Chen, W., Vermaak, I., Viljoen, A., Non-destructive quality assessment of herbal tea blends using hyperspectral imaging, *Phytochem. Lett.* 24 (2018) 94-101.  
<https://doi.org/10.1016/j.phytol.2018.01.016>.
- [60] Kimuli, D., Wang, W., Wang, W., Jiang, H., Zhao, X., Chu, X., Application of SWIR hyperspectral imaging and chemometrics for identification of aflatoxin B1 contaminated maize kernels, *Infrared Phys. Technol.* 89 (2018) 351-362.  
<https://doi.org/10.1016/j.infrared.2018.01.026>.
- [61] Cortés, V., Cubero, S., Aleixos, N., Blasco, J., Talens, P., Sweet and nonsweet taste discrimination of nectarines using visible and near-infrared spectroscopy, *Postharvest Biol. Technol.* 133 (2017) 113-120.  
<https://doi.org/10.1016/j.postharvbio.2017.07.015>.
- [62] Shrestha, S., Knapic, M., Zibrat, U., Deleuran, L. C., Gislum, R., Single seed near-infrared hyperspectral imaging in determining tomato (*Solanum lycopersicum* L.) seed quality in association with multivariate data analysis, *Sens. Actuators B Chem.* 237 (2016) 1027-1034. <https://doi.org/10.1016/j.snb.2016.08.170>.
- [63] Santos, P. M., Pereira-Filho, E. R., Colnago, L. A. Detection and quantification of milk adulteration using time domain nuclear magnetic resonance (TD-NMR), *Microchem. J.* 124 (2016) 15-19.  
<https://doi.org/10.1016/j.microc.2015.07.013>.
- [64] Chen, Z., Gu, A., Zhang, X., Zhang, Z., Authentication and inference of seal stamps on Chinese traditional painting by using multivariate classification and near-infrared spectroscopy, *Chemom. Intell. Lab. Syst.* 171 (2017) 226-233.  
<https://doi.org/10.1016/j.chemolab.2017.10.017>.
- [65] Oliveira, V. S., Honorato, R. S., Honorato, F. A., Pereira, C. F., Authenticity assessment of banknotes using portable near infrared spectrometer and

- chemometrics, *Forensic Sci. Int.* 286 (2018) 121-127. <https://doi.org/10.1016/j.forsciint.2018.03.001>.
- [66] Martins, A. R., Talhavini, M., Vieira, M. L., Zacca, J. J., Braga, J. W. B., Discrimination of whisky brands and counterfeit identification by UV-Vis spectroscopy and multivariate data analysis, *Food Chem.* 229 (2017) 142-151. <https://doi.org/10.1016/j.foodchem.2017.02.024>.
- [67] Costa, V. C., Aquino, F. W. B., Paranhos, C. M., Pereira-Filho, E. R., Identification and classification of polymer e-waste using laser-induced breakdown spectroscopy (LIBS) and chemometric tools, *Polym. Test.* 59 (2017) 390-395. <https://doi.org/10.1016/j.polymertesting.2017.02.017>.
- [68] Silva, N. C., Pimentel, M. F., Honorato, R. S., Talhavini, M., Maldaner, A. O., Honorato, F. A., Classification of Brazilian and foreign gasolines adulterated with alcohol using infrared spectroscopy, *Forensic Sci. Int.* 253 (2015) 33-42. <https://doi.org/10.1016/j.forsciint.2015.05.011>.
- [69] Silva, M. P. F., Rodrigues e Brito, L., Honorato, F. A., Paim, A. P. S., Pasquini, C., Pimentel, M. F., Classification of gasoline as with or without dispersant and detergent additives using infrared spectroscopy and multivariate classification, *Fuel* 116 (2014) 151-157. <https://doi.org/10.1016/j.fuel.2013.07.110>.
- [70] Sinkov, N. A., Sandercock, P. M. L., Harynyuk, J. J., Chemometric classification of casework arson samples based on gasoline content, *Forensic Sci. Int.* 235 (2014) 24-31. <https://doi.org/10.1016/j.forsciint.2013.11.014>.
- [71] Kovalishyn, V., Poda, G., Efficient variable selection batch pruning algorithm for artificial neural networks, *Chem. Intell. Lab. Syst.* 149 (2015) 10-16. <https://doi.org/10.1016/j.chemolab.2015.10.005>.
- [72] Li, S., Yang, B., Qi, F., Accelerate global sensitivity analysis using artificial neural network algorithm: Case studies for combustion kinetic model, *Combust. Flame* 168 (2016) 53-64. <https://doi.org/10.1016/j.combustflame.2016.03.028>.
- [73] Bian, X. H., Li, S. J., Fan, M. R., Guo, Y. G., Chang, N., Wang, J. J., Spectral quantitative analysis of complex samples based on the extreme learning machine, *Anal. Methods* 8 (23) (2016) 4674-4679. <https://doi.org/10.1039/C6AY00731G>.
- [74] Viejo, C. G., Fuentes, S., Torrico, D., Howell, K., Dunshea, F. R., Assessment of beer quality based on foamability and chemical composition using computer vision algorithms, near infrared spectroscopy and machine learning algorithms, *J. Sci. Food Agr.* 9 8 (2) (2018) 618-627. <https://doi.org/10.1002/jsfa.8506>.
- [75] Oliveira, H. C., Cunha, J. C. E., Rocha, J. C., Nunez, E. G. F., Rapid monitoring of beer-quality attributes based on UV-Vis spectral data, *Int. J. Food Prop.* 20 (2017) 1686-1699. <https://doi.org/10.1080/10942912.2017.1352602>.
- [76] Das, B., Sahoo, R. N., Pargal, S., Krishna, G., Verma, R., Chinnusamy, V., Sehgal, V. K., Gupta, V. K., Dash, S. K., Swain, P., Quantitative monitoring of sucrose, reducing sugar and total sugar dynamics for phenotyping of water-deficit stress tolerance in rice through spectroscopy and chemometrics, *Spectrochim. Acta Part A* 192 (2018) 41-51. <https://doi.org/10.1016/j.saa.2017.10.076>.
- [77] Li, Z., Zhang, X., Karanassios, V., Verification of the performance of artificial neural networks (ANNs) versus partial least squares (PLS) for spectral interference correction in optical emission spectrometry, *Proceedings of SPIE 9118, Independent Component Analyses, Compressive Sampling, Wavelets, Neural Net, Biosystems, and Nanoengineering XII*, 911812 (2014). <https://doi.org/10.1117/12.2050326>.
- [78] Li, Z., Zhang, X., Karanassios, V., How do artificial neural networks (ANNs) compare to partial least squares (PLS) for spectral interference correction in optical emission spectrometry? *Proc. SPIE 9496, Independent Component Analyses, Compressive Sampling, Large Data Analyses (LDA), Neural Networks, Biosystems, and Nanoengineering XIII*, 94960M (2015). <https://doi.org/10.1117/12.2177516>.
- [79] Li, H. H., Zhang, B., Hu, W. W., Liu, Y., Dong, C. W., Chen, Q. S., Monitoring black tea fermentation using a colorimetric sensor array-based artificial olfaction system, *J. Food Process. Pres.* 42 (1) (2018). <https://doi.org/10.1111/jfpp.13348>.
- [80] Dan, S. J., Yang, S. X., Tian, F. C., Den, L., Classification of orange growing locations based on the near-infrared spectroscopy using data mining, *Intell. Autom. Soft Co.* 22 (2) (2016) 229-236. <https://doi.org/10.1080/10798587.2015.1095474>.
- [81] Wang, J. M., Liao, X. Y., Zheng, P. C., Xue, S. W., Peng, R., Classification of Chinese herbal medicine by laser-induced breakdown spectroscopy with principal component analysis and artificial neural network, *Anal. Lett.* 51(4) (2018) 575-586. <https://doi.org/10.1080/00032719.2017.1340949>.
- [82] Wang, F., Wang, B., Wang, L., Xiong, Z. Y., Gao, W., Li, P., Li, H. J., Discovery of discriminatory quality control markers for Chinese herbal medicines and

related processed products by combination of chromatographic analysis and chemometrics methods: *Radix Scutellariae* as a case study, *J. Pharmaceut. Biomed.* 138 (2017) 70-79. <https://doi.org/10.1016/j.jpba.2017.02.004>.

[83] Ding, G. Y., Wang, Y. S., Liu, A. N., Hou, Y. Y., Zhang, T. J., Bai, G., Liu, C. X., From chemical markers to quality markers: an integrated approach of UPLC/Q-TOF, NIRS, and chemometrics for the quality assessment of honeysuckle buds, *RSC Adv.* 7 (36) (2017) 22034-22044. <https://doi.org/10.1039/C6RA28152D>.

[84] Ito, S., Barchi, A. C., Escaramboni, B., Neto, P. D., Herculano, R. D., Borges, F. A., Miranda, M. C. R., Nunez, E. G. F., UV/Vis spectroscopy combined with chemometrics for monitoring solid-state fermentation with *Rhizopus microsporus* var. *oligosporus*, *J. Chem. Technol. Biot.* 92 (10) (2017) 2563-2572. <https://doi.org/10.1002/jctb.5271>.

[85] Li, K. H., Guo, L. B., Li, J. M., Yang, X. Y., Yi, R. X., Li, X. Y., Lu, Y. F., Zeng, X. Y., Quantitative analysis of steel samples using laser-induced breakdown spectroscopy with an artificial neural network incorporating a genetic algorithm, *Appl. Optics* 56 (4) (2017) 935-941. <https://doi.org/10.1364/AO.56.000935>.

[86] Zhang, P., Sun, L. X., Kong, H. Y., Yu, H. B., Guo, M. T., Zeng, P., A method derived from genetic algorithm, principal component analysis and artificial neural networks to enhance classification capability of laser-induced breakdown spectroscopy, *Proc. SPIE* 10461, AOPC 2017: Optical Spectroscopy and Imaging, 1046107 (24 October 2017). <https://doi.org/10.1117/12.2281493>.

[87] Gurbanov, R., Bilgin, M., Severcan, F., Restoring effect of selenium on the molecular content, structure and fluidity of diabetic rat kidney brush border cell membrane, *Biochim. Biophys. Acta – Biomembr.* 1858 (4) (2016) 845-854. <https://doi.org/10.1016/j.bbmem.2016.02.001>.

[88] Akbari Hasanjani, H. R., Sohrabi, M. R., Artificial neural networks (ANN) for the simultaneous spectrophotometric determination of fluoxetine and sertraline in pharmaceutical formulations and biological fluid, *Iran J. Pharm. Res.* 16 (2) (2017) 478-489.

[89] Guo, Y., Ni, Y. N., Chen, J. F., Kokot, S., A kinetic spectrofluorometric method, aided by chemometrics, for the analysis of sibutramine, indapamide and hydrochlorothiazide compounds found in weight-reducing tonic samples, *Anal. Methods* 8 (1) (2016) 197-204. <https://doi.org/10.1039/C5AY02191J>.

[90] Leardi, R., Experimental design in chemistry: A tutorial, *Anal. Chim. Acta* 652 (2009) 161-172. <https://doi.org/10.1016/j.aca.2009.06.015>.

[91] Barros Neto, B., Scarminio, I. S., Bruns, R. E. Como fazer experimentos, Bookman: Porto Alegre, 2010.

[92] Pereira, F. M. V., Pereira-Filho, E. R., Aplicação de programa computacional livre em planejamento de experimentos: um tutorial, *Quim. Nova* 41 (2018) 1061-1071. <https://doi.org/10.21577/0100-4042.20170254>.

[93] Pereira Filho, E. R., Planejamento fatorial em química: maximizando a obtenção de resultados. Edufscar: São Carlos, 2015.

[94] Ferreira, S. L. C., Santos, W. N. L., Quintella, C. M., Barros Neto, B., Bosque-Sendra, J. M., Doehlert matrix: a chemometric tool for analytical chemistry – review, *Talanta* 63 (2004) 1061-1067. <https://doi.org/10.1016/j.talanta.2004.01.015>.

[95] Ferreira, S. L. C., Bruns, R. E., Ferreira, H. S., Matos, G. D., David, J. M., Brandão, G. C., da Silva, E. G. P., Portugal, L. A., dos Reis, P. S., Souza, A. S., dos Santos, W. N. L., Box-Behnken desing: an alternative for the optimization of analytical methods, *Anal. Chim. Acta* 597 (2007) 179-186. <https://doi.org/10.1016/j.aca.2007.07.011>.

[96] Ferreira, S. L. C., Introdução às técnicas de planejamento de experimentos. Vento Leste: Salvador, 2015.

[97] Teófilo, R. F., Ferreira, M. M. C., Quimiometria II: planilhas eletrônicas para cálculos de planejamentos experimentais, um tutorial, *Quim. Nova.* 29 (2006) 338-350. [http://quimicanova.sbq.org.br/imagebank/pdf/Vol29No2\\_338\\_25-DV04378.pdf](http://quimicanova.sbq.org.br/imagebank/pdf/Vol29No2_338_25-DV04378.pdf).

[98] Myers, R. H., Montgomery, D. C., Anderson-Cook, C. M., Response surface methodology: process and product optimization using designed experiments, Wiley: Hoboken, 2009.

[99] Hewavitharana, A. K., Abu Kassim, N. S., Shaw, P. N., Standard addition with internal standardization as an alternative to using stable isotope labelled internal standard to correct for matrix effects-comparison and validation using liquid chromatography – tandem mass spectrometric assay of vitamin D, *J. Chromatogr. A* 1553 (2018) 101-107. <https://doi.org/10.1016/j.chroma.2018.04.026>.

- [100] Andri, B., Dispas, A., Klinkenberg, R., Streeb, B., Marini, R. D., Ziemons, E., Hubert, P., Is supercritical fluid chromatography hyphenated to mass spectrometry suitable for the quality control of vitamin D3 oily formulations? *J. Chromatogr. A* 1515 (2017) 209-217. <https://doi.org/10.1016/j.chroma.2017.07.057>.
- [101] Finetti, C., Sola, L., Elliott, J., Chiari, M., Synthesis of hydrogel via click chemistry for DNA electrophoresis, *J. Chromatogr. A* 1513 (2017) 226-234. <https://doi.org/10.1016/j.chroma.2017.07.042>.
- [102] Doriva-Garcia, N., Bones, J., Monitoring leachables from single-use bioreactor bags for mammalian cell culture by dispersive liquid-liquid microextraction followed by ultra high performance liquid chromatography quadrupole time of flight mass spectrometry. *J. Chromatogr. A* 1512 (2017) 51-60. <https://doi.org/10.1016/j.chroma.2017.06.077>.
- [103] Zimmermann, A., Totoli, E. G., Fernandes, F. H. A., Salgado, H. R. N., An eco-friendly and low-cost method for the quantification of cefazolin sodium in powder for injectable solution using thin-layer chromatography assisted by digital images, *JPC-J. Planar Chromat.* 30 (2017) 285-290. <https://doi.org/10.1556/1006.2017.30.4.8>.
- [104] Risoluti, R., Materazzi, S., MicroNIR/Chemometrics assesment of occupational exposure to hydroxyurea, *Front. Chem.* 6 (2018) 1-9. <https://doi.org/10.3389/fchem.2018.00228>.
- [105] Dunst, K. J., Trzcinski, K., Scheibe, B., Study of the NO<sub>2</sub> sensing mechanism of PEDOT-RGO film using *in situ* Raman spectroscopy, *Sens. Actuator B-Chem.* 260 (2018) 1025-1033. <https://doi.org/10.1016/j.snb.2018.01.089>.
- [106] Hogmalm, K. J., Zack, T., Karlsson, A. K. O., Sjoqvist, A. S. L., Garbe-Schonberg, D., *In situ* Rb-Sr and K-Ca dating by LA-ICP-MS/MS: an evaluation of N<sub>2</sub>O and SF<sub>6</sub> as reaction gases, *J. Anal. At. Spectrom.* 32 (2017) 305-313. <https://doi.org/10.1039/C6JA00362A>.
- [107] Fellhauer, D., Gaona, X., Rothe, J., Altmaier, M., Fanghanel, T., Neptunium(VI) solubility in alkaline CaCl<sub>2</sub> solutions: evidence for the formation of calcium neptunates Ca<sub>x</sub> NpO<sub>3+x</sub> (s,hyd), *Monatsh. Chem.* 149 (2018) 237-252. <https://doi.org/10.1007/s00706-017-2116-4>.
- [108] Wang, X., He, H., Fan, X. G., Tang, M., Signal processing method for Raman spectra based on matching pursuit, *Spectrosc. Spect. Anal.* 38 (2018) 93-98. [https://doi.org/10.3964/j.issn.1000-0593\(2018\)01-0093-06](https://doi.org/10.3964/j.issn.1000-0593(2018)01-0093-06).
- [109] Liu, T., Zhang, W. J., Zhang, Z., Chen, M. L., Wang, J. H., Qian, X. H., Qin, W. J., Sensitive western-blot analysis of azide-tagged protein post translation modifications using thermoresponsive polymer self-assembly, *Anal. Chem.* 90 (2018) 2186-2192. <https://doi.org/10.1021/acs.analchem.7b04531>.
- [110] Orihara, K., Hikichi, A., Arita, T., Muguruma, H., Yoshimi, Y., Heparin molecularly imprinted Polymer thin film on gold electrode by plasma-induced graft polymerization for label-free biosensor, *J. Pharmaceut. Biomed.* 151 (2018) 324-330. <https://doi.org/10.1016/j.jpba.2018.01.012>.
- [111] Cinti, S., Mazzaracchio, V., Orturk, G., Moscone, D., Arduini, F., A lab-on-a-tip approach to make electroanalysis user-friendly and decentralized: detection of copper ions in river water, *Anal. Chim. Acta* 1029 (2018) 1-7. <https://doi.org/10.1016/j.aca.2018.04.065>.
- [112] Coutinho, M. S., Morais, C. L. M., Neves, A. C. O., Menezes, F. G., Lima, K. M. G., J., Colorimetric determination of ascorbic acid based on its interfering effect in the enzymatic analysis of glucose: an approach using smartphone image analysis, *J. Braz. Chem. Soc.* 28 (2017) 2500-2505. <https://doi.org/10.21577/0103-5053.20170086>.
- [113] Sandusky, P. O., Introducing undergraduate students to metabolomics using a NMR-based analysis of coffee beans, *J. Chem. Educ.* 94 (2017) 1324-1328. <https://doi.org/10.1021/acs.jchemed.6b00559>.
- [114] Ayed, L., Bakir, K., Ben Mansour, H., Hammami, S., Cheref, A., Bakhrouf, A., *In vitro* mutagenicity, NMR metabolite characterization of azo and triphenylmethanes dyes by adherents bacteria and the role of the “can” adhesion gene in activated sludge, *Microb Pathog* 103 (2017) 29-39. <https://doi.org/10.1016/j.micpath.2016.12.016>.