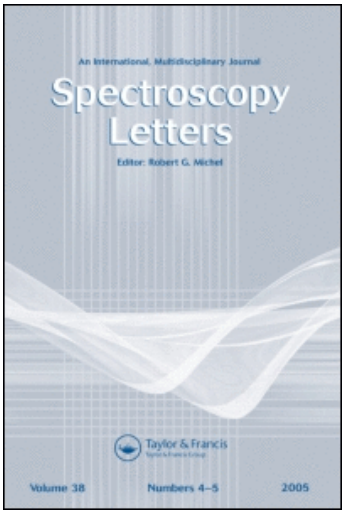

Spectroscopy Letters
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597299

Quantitative Infrared Spectroscopy in the Undergraduate Laboratory Via Multivariate Mixture Analysis of a Simulated Analgesic

Scot W. Huffman[a]; Arthur L. Salido[a]; David D. Evanoff Jr.[a]
[a] Department of Chemistry and Physics, Western Carolina University, Cullowhee, NC

Online publication date: 20 October 2010

To cite this Article Huffman, Scot W. , Salido, Arthur L. and Evanoff Jr., David D.(2010) 'Quantitative Infrared Spectroscopy in the Undergraduate Laboratory Via Multivariate Mixture Analysis of a Simulated Analgesic', Spectroscopy Letters, 43: 7, 539 — 544
To link to this Article: DOI: 10.1080/00387010.2010.510702
URL: http://dx.doi.org/10.1080/00387010.2010.510702

Spectroscopy Letters, 43:539–544, 2010

ISSN: 0038-7010 print/1532-2289 online
DOI: 10.1080/00387010.2010.510702

Taylor & Francis
Taylor & Francis Group

Quantitative Infrared Spectroscopy in the Undergraduate Laboratory Via Multivariate Mixture Analysis of a Simulated Analgesic

Scot W. Huffman, Arthur L. Salido, and David D. Evanoff, Jr.

Department of Chemistry and Physics, Western Carolina University, Cullowhee, NC

ABSTRACT Infrared spectroscopy has been utilized for nearly a century as a qualitative tool for, at least, partial structure elucidation. More recently, infrared spectroscopy has become a powerful, reliable quantitative analysis tool in industrial chemistry applications, especially in the areas of quality control. Unfortunately, this newer aspect of infrared spectroscopy is often not conveyed during the undergraduate chemical education experiences. We describe an undergraduate laboratory experiment that illustrates the application of infrared spectroscopy combined with multivariate data analysis in the concentration determination of a three-component solid state mixture.

KEYWORDS analgesic, infrared spectroscopy, multivariate

BACKGROUND AND MOTIVATION

Although infrared spectroscopy can be a powerful tool for quantitative analysis, undergraduate curricula tend to teach infrared spectroscopy as a qualitative technique, an approach that our department has historically followed. To correct this oversight, we developed the laboratory experiment, described herein. In our chemistry curriculum, we have a three-course lecture–lab sequence in the analytical chemistry division, of which every student is required to take the first two courses, and of which a subset of students takes the third course. In the first course, Quantitative Analysis, we cover equilibrium chemistry and electrochemistry via volumetric and gravimetric analysis, spectroscopy (UV/Vis), and potentiometry. In the second course, Instrumental Analysis I, we cover briefly instrumental design, method development, and data analysis of the instrumentation typical in chemical laboratories in industry and academia including but not limited to ultraviolet/visible (UV/Vis), Fourier transform infrared (FT-IR), flame atomic absorption (FAAS), and fluorescence spectroscopies. In the third course, Instrumental Analysis II, we cover instrumental design in detail and cover method development and data analysis of instrumentation that is less typical or found at research facilities, including but not limited to

Received 1 September 2009; accepted 26 October 2009.

Address correspondence to Scot W. Huffman, Department of Chemistry and Physics, 1 University Drive, Western Carolina University, Cullowhee, NC 28723. E-mail: shuffman@wcu.edu

Raman and infrared spectroscopy and spectroscopic imaging, SEM, TEM, AFM, STM, CE, cavity ring down spectroscopy, LC-MS, and multidimensional NMR. The laboratory experiment described herein is taught in the middle course, Instrumental Analysis I, and is designed for junior-level students. This experimental exercise is performed by the students in small groups as part of a round-robin system of experiments that encompasses the second half of the semester. In this round-robin system, each experiment lasts for two 3 hr laboratory periods, and the report is due 1 week after the second laboratory period.

In the literature, there are many excellent examples of the teaching of statistics and in particular of multivariate data analysis in chemistry at the undergraduate and graduate levels. These examples range from philosophy of teaching the data analysis aspects of analytical chemistry,[1–3] to tutorials on data analysis,[4,5] to descriptions of laboratory exercises involving multivariate data analysis.[6–11] Herein, we present a description of a laboratory exercise where the students utilize infrared spectroscopy and multivariate data analysis to preform quantitative characterization of the data. Because infrared spectroscopies, both near and mid are becoming so prevalent in the process analytical world,[12–14] we have chosen the following setting for the experiment. The students read in the laboratory handout that they work as a process analytical chemist for a pharmaceutical company. Their current assignment is to develop a method for validating the concentration of an analgesic in a powder formulation of three components. As for reporting their work, the students must write a standard operating procedure to be followed by technicians.

LEARNING OBJECTIVES

We have designed this laboratory experiment to achieve several learning objectives that reinforce the general laboratory theme of method development. These objectives are the effects of spectral overlap on fitting data, the use of linear algebra in fitting multivariate data, the validation and statistical evaluation of calibration models, and the usefulness of infrared spectroscopy in quantitative analysis.[12–14]

MATERIALS AND METHODS

In this experiment, the students analyzed an analgesic powder mixture using a single bounce diamond and ZnSe attenuated total reflectance (ATR) sampling accessory attached to a PerkinElmer Spectrum One FT-IR (Waltham, MA). The spectral parameters for all measurements were $4\,cm^{-1}$ resolution, weak apodization, four coadded scans for both the background and sample, data spacing of $1\,cm^{-1}$, magnitude-based phase correction, and 0.20 cm/s interferometer speed. Similar measurements could be made with a diffuse reflectance accessory if that accessory were available instead.

Creation of Calibration Standards

Initially the students created five calibration standards containing dextrose (Fisher Scientific), stearic acid (Kanto Chemical Company), and acetaminophen (Mallinckrodt) in the ratios listed in Table 1. To prevent excessive waste, the total mass of the mixtures did not exceed 0.1 g. The mixtures were homogenized with a Wig-L-Bug (International Crystal Laboratories) for 30s, and stored in 1–2-mL capped glass vials (Fisher Scientific) for later measurement.

Measurement of Calibration Standards

The ATR was cleaned before every measurement using water and then acetone. The background single-beam spectra were acquired using a clean and empty ATR and reacquired after every five sample measurements. In order to ensure accurate sampling of the solid mixture we took five aliquots from each standard mixture and measured. Also each pure component of the mixture was measured using the same measurement parameters as the background and mixture samples; these pure component data were subsequently used to select analytical wave numbers.

Measurement of Validation Mixtures

The validation sample was a mixture of the three components, whose concentration had been determined by other means by the instructor. The infrared

TABLE 1 Mixture Parameters

Component	% by weight
Dextrose (filler)	60–95
Stearic acid (binder)	4–15
Acetaminophen (active ingredient)	4–25

spectroscopic measurements of the validation sample were made using the same procedures as these for the calibration standards.

RESULTS AND DISCUSSION

All of the spectra were saved as ASCII-formatted text files with the wavenumber and absorbance data in two columns separated by a tab. These data were then loaded into either Excel (Microsoft) or Octave[15] (a free, open-source matrix math program).

Selection of Analytical Wavenumbers

Although whole spectra can be used in these calculations, often with better fitting results,[16,17] we have chosen to use only one absorption maximum per mixture component. This choice is motivated by the students' lack of experience in matrix algebra and the software used for these computations; i.e., it is conceptually simpler for the students to understand one wavenumber per component.

The students were instructed to select a wave number that corresponds to an absorption maximum of each pure component, but because of the variable, high humidity of the measurement environment, the spectral range is limited to regions that fall outside regions that show heavy effects from water vapor (approximately 700–1350 cm^{-1}). This was typically done by manual, visual inspection of all three pure component spectra in either overlapped or stacked spectral arrangement. In any graphing program, the stacked spectral arrangement can be accomplish by adding a different constant to the absorbance values of the three pure component spectra. During this selection process the students found that, with the exception of the carbonyl band of the stearic acid, there are very few bands that do not at least partially overlap with bands from other components as can be seen in the spectra shown in Fig. 1. Consequently, they were instructed to do their best. As part of their report, the students assigned the vibrational mode for the analytical wavenumbers that they selected. These students, whose data is shown herein, selected 994, 1297, and 1225 cm^{-1} and assigned these bands to the C-O stretch of dextrose, the C-H deformation mode of the aliphatic chain of the stearic acid, and the C-H in-plane rocking mode on the aromatic ring of acetaminophen, respectively.

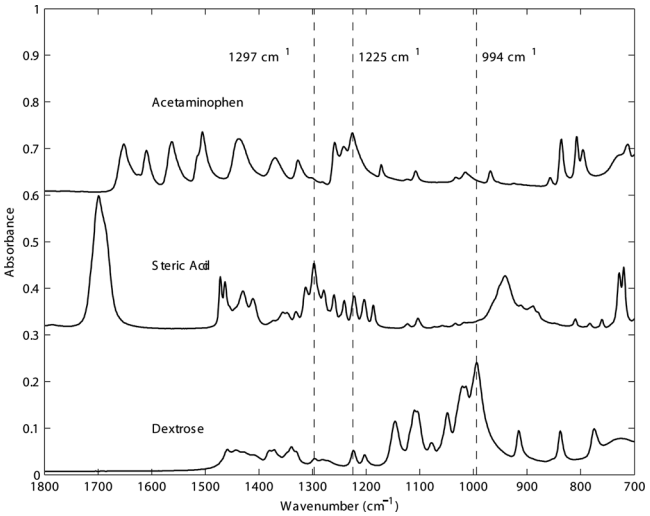

FIGURE 1 **ATR-FT-IR spectra of neat dextrose, stearic acid, and acetaminophen. The vertical dashed lines indicate the analytical wavenumbers used during data analysis.**

Averaging Spectral Data

The students averaged all five repetitions of their spectra. The average spectra of the calibration set are shown in Fig. 2. The averaged absorbance values at three analytical wavenumbers were extracted from the whole spectra of calibration and validation samples and assembled into a matrix such as the ones presented in Tables 2 and 3 populated with student data.

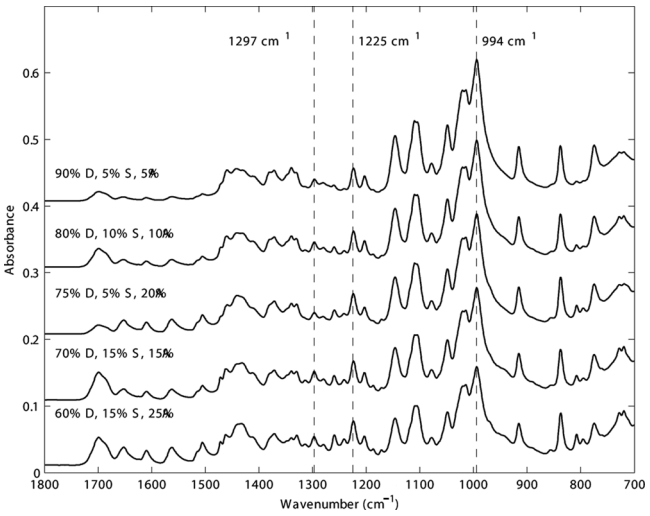

FIGURE 2 **ATR-FT-IR average spectra of the five mixtures in the calibration set. The vertical dashed lines indicate the analytical wavenumbers used during data analysis.**

TABLE 2 Student Calibration Standard Average Spectral Data

$A_{\bar{\nu}=994\text{cm}^{-1}}$	$A_{\bar{\nu}=1297\text{cm}^{-1}}$	$A_{\bar{\nu}=1225\text{cm}^{-1}}$
0.15907	0.054111	0.076502
0.17772	0.051848	0.066005
0.18859	0.04005	0.068032
0.19902	0.046125	0.061019
0.22008	0.04017	0.055802

Univariate Model

When spectral features are free of interferences, a straightforward univariate analysis of Beer's Law can be performed. In previous courses, students have utilized the classic method of least squares to determine a linear equation that best fits their data using Excel built-in functions. Consequently, this portion of this laboratory exercise is not a new experience, but is necessary for understanding the limitation of fitting data with overlapping spectral features or interferences, such as those seen in Fig. 1. For the acetaminophen absorbance presented in Table 2 for wavenumber 1225 cm^{-1}, the fitting parameters are the slope, 0.00096823 cm, and the y-intercept, 0.050949. The plot of absorbance versus concentration is shown in Fig. 3.

Multivariate Model

The first laboratory exercise for the course is entitled, "Introduction to Data Processing, Analysis and Visualization." In it the students learned, among other things, basic linear algebra, and actually did this regression analysis without the chemical context but with instructor-supplied data. So this is not the first time they were required to perform the task. For completeness, we will describe this procedure herein.

For each calibration standard in a calibration set, the absorbance values, a, are equated to the concentrations, c, of individual components (x, y, or z) as shown in these three equations,

$$\begin{aligned} a_m &= k_{xm}c_x + k_{ym}c_y + k_{zm}c_z \\ a_n &= k_{xn}c_x + k_{yn}c_y + k_{zn}c_z \\ a_o &= k_{xo}c_x + k_{yo}c_y + k_{zo}c_z \end{aligned} \quad (1)$$

where the indices m, n, and o indicate the three different wavenumbers, 994, 1297, and 1225 cm^{-1}, respectively. The coefficients, k, in these equations are the proportion that each component linearly adds to the overall absorbance at the particular wavenumber (m, n, or o).

These equations can be written in matrix form as shown in Eq. (2):

$$\begin{pmatrix} a_m & a_n & a_o \end{pmatrix} = \begin{pmatrix} c_x & c_y & c_z \end{pmatrix} \begin{pmatrix} k_{xm} & k_{ym} & k_{zm} \\ k_{xn} & k_{yn} & k_{zn} \\ k_{xo} & k_{yo} & k_{zo} \end{pmatrix} \quad (2)$$

Then, these matrices can be expanded to include all five mixtures in the calibration set as indicated by the numerical index in the matrices in Eq. (3):

$$\begin{pmatrix} a_{1m} & a_{1n} & a_{1o} \\ a_{2m} & a_{2n} & a_{2o} \\ a_{3m} & a_{3n} & a_{3o} \\ a_{4m} & a_{4n} & a_{4o} \\ a_{5m} & a_{5n} & a_{5o} \end{pmatrix} = \begin{pmatrix} c_{1x} & c_{1y} & c_{1z} \\ c_{2x} & c_{2y} & c_{2z} \\ c_{3x} & c_{3y} & c_{3z} \\ c_{4x} & c_{4y} & c_{4z} \\ c_{5x} & c_{5y} & c_z \end{pmatrix} \begin{pmatrix} k_{xm} & k_{ym} & k_{zm} \\ k_{xn} & k_{yn} & k_{zn} \\ k_{xo} & k_{yo} & k_{zo} \end{pmatrix} \quad (3)$$

Equation (3) can be written in a more compact manner as shown in Eq. (4),

$$\mathbf{A} = \mathbf{CK} \quad (4)$$

where $\mathbf{A}$ is a matrix of absorbance values corresponding to the analytical wave numbers (columns) for each mixture (rows), $\mathbf{C}$ is the concentration matrix of each component (columns) in each mixture (rows), and $\mathbf{K}$ is the regression matrix that is the best fitting of coefficients that linearly relates the

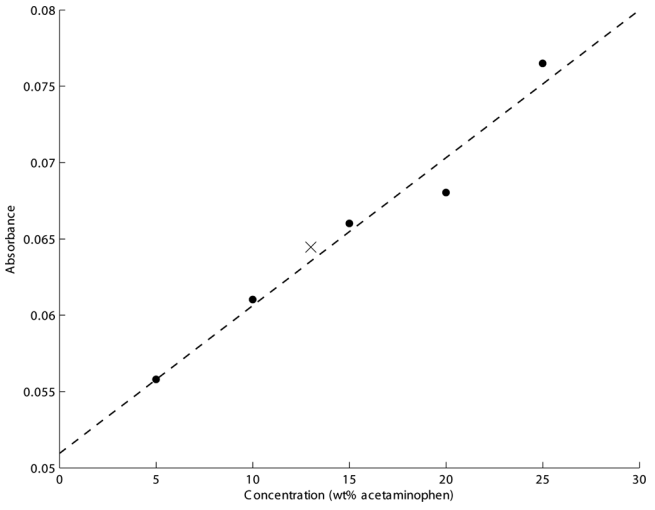

FIGURE 3 Univariate Beer's Law graph of the absorbance at 1225 cm^{-1}. The straight dashed line is the best of fit line for the standard calibration data. The "×" indicates the position of the absorbance and the $c_{predicted}$ for the validation sample.

absorbance to the concentration for each component (columns) in every mixture (rows).

Fitting the calibration standards using all the components in a mixture involves solving Eq. (4) for **K**. If **C** is a square matrix this operation can be accomplished by left-multiplying both sides of the equation by the inverse of **C**, $\mathbf{C}^{-1}$. In this experiment, however, the matrix has dimensions of five mixtures by three components. Therefore, the pseudo-inverse method was used, resulting in Eq. (5).[18]

$$\mathbf{K} = (\mathbf{C}^T\mathbf{C})^{-1}\mathbf{C}^T\mathbf{A} \tag{5}$$

where $\mathbf{C}^T$ is the transpose of **C**. These matrix operations are relatively straightforeward in Excel and Octave. For the data presented in Table 2, the best fit model is the **K** matrix shown in Eq. (6):

$$\mathbf{K} = \begin{matrix} 0.0024004 & 0.00033107 & 0.00049477 \\ 0.00032572 & 0.0015696 & 0.00076244 \\ 0.00037364 & 0.00039567 & 0.0013827 \end{matrix} \tag{6}$$

Prediction of Validation Sample Concentrations

After the data has been fit, the students used both their univariate and multivariate models or fitting parameters to predict the concentrations of the acetaminophen of the validation sample. For the single univariate model, the students solved for the concentration of acetaminophen using their univariate fitting parameters as shown in Eq. (7),

$$C_{acetaminophen} = \frac{\text{Absorbance}_{1225\text{cm}^{-1}} - \text{slope}}{\text{y} - \text{intercept}} \tag{7}$$

which results in the concentration of 13.9 wt% using the data listed in Tables 2 and 3. For most students, it was immediately clear that something was wrong, as shown in Fig. 3 as a scatter plot of the absorbance at $1225\,\text{cm}^{-1}$ versus concentration of acetaminophen, where the 20-wt% absorbance appears low. One temptation that the students had was to drop the 20 wt%. They were instructed to wait and see what the multivariate model would produce.

Predicting the concentration of acetaminophen in the validation sample with the multivariate model was accomplished by using the pseudo-inverse solution for the concentration shown in Eq. (8).

$$\mathbf{C}_{predicted} = \mathbf{A}_{validation}\mathbf{K}^T(\mathbf{K}\mathbf{K}^T)^{-1} \tag{8}$$

where $\mathbf{C}_{predicted}$ is the matrix of predicted concentrations of the three components, and $\mathbf{A}_{validation}$ is the measured absorbance values of the validation mixture. This calculation yielded the following results: 77.4 wt% dextrose, 3.6 wt% stearic acid, and 16.9 wt% acetaminophen. This validation standard was issued to the students with known concentrations of 78.0 wt% dextrose, 4.5 wt% stearic acid, and 17.6 wt% acetaminophen. So compared to the single absorbance value fitting of the data, this method provides a closer prediction.

TABLE 3 Student Validation Sample Average Spectral Data

$A_{\bar{\nu}=994\text{cm}^{-1}}$	$A_{\bar{\nu}=1297\text{cm}^{-1}}$	$A_{\bar{\nu}=1225\text{cm}^{-1}}$
0.19337	0.037942	0.06445

Comparison of Models

The students were asked to compare the univariate method with the multivariate method of modeling their data. For this task, they compared the known validation concentrations with their predicted concentrations in both models using the *t* test.[19] In this test a 95% confidence interval, μ, is calculated using Eq. (9),

$$\mu = \bar{c}_{predicted} \pm \frac{ts}{\sqrt{N}} \tag{9}$$

where $\bar{c}_{predicted}$ is the average predicted concentration, N is the number of measurements, t is the t statistic for a 95% confidence interval, and s is the standard deviation across the five measurements. If their data is statistically the same as the known validation concentration, the known concentration should fall within the interval. The intervals are 16.9 ± 0.8 wt% for the multivariate model and 13 ± 2 wt% for the univariate model. Although, for this set of data, neither model provides a perfect fit, there is a 95% chance that the true concentration lies within the confidence interval of the multivariate model. In contrast, for the univariate model, the known concentration does not lie within the confidence interval.

The students were asked to provide an explanation for this difference in performance. An acceptable explanation of the poorer behavior of the univariate model is that the dextrose, stearic acid, and acetaminophen all contribute to the height of the apparent "acetaminophen" band at 1225 cm^{-1}, and the concentrations of these three components are changing across the samples at different rates and in different directions. This complicated behavior results in apparently "noisy" data in a Beer's law plot as shown in Fig. 3. In contrast, the multivariate model works much better because it accounts for all components in the mixture that contribute to the absorbance values.

CONCLUSIONS

This comparison illustrate the limitations of the classic least squares model, which is that all contributing components to the region of spectral interest must be accounted for in the model. Although not illustrated here, this limitation is true for both the univariate and multivariate methods. If, for example, the multivariate model had only modeled the acetaminophen and stearic acid analytical wave numbers, then it too would not have produced a statistically accurate result.

In this experiment the students sharpened their basic linear algebra skills, learned a method for and the importance of compensating for spectral interferences (which can be considered matrix effects), and reinforced their knowledge of statistically comparing two methods of analysis. They also learned that infrared spectroscopy can be used for quantitative analysis. We suggest that additions and changes could be made to this laboratory experiment to illustrate other aspects of method development. These might include the determination of the optimum number of aliquots to measure that one should illustrate the need for generating enough samples to adequately represent the population; and an increase in the number of analytical wavenumbers from two whole or nearly-whole spectra from three whole or nearly–whole spectra, to illustrate further the benefits of multivariate mixture analysis.

REFERENCES

1. Brown, S. D. Has the chemometrics revolution ended? some views on the past, present and future of chemometrics. *Chemometr. Intell. Lab.* **1995**, *30*(1), 1–15.
2. Msimanga, H. Z.; Elkins, P.; Tata, S. K.; Smith, D. R. A chemometics module for an undergraduate instrumental analysis chemistry course. *J. Chem. Educ.* **2005**, *82*(3), 415–422.
3. Wegscheider, W. Standardization, quality control and education in analytical chemistry. *Fresen. J. Anal. Chem.* **1994**, *349*(12), 784–793.
4. Melucci, D. Manual data processing in analytical chemistry: Linear calibration. *J. Chem. Educ.* **2008**, *85*(10), 1346.
5. Brereton, R. G. *Applied Chemometrics for Scientists*; John Wiley & Sons, Ltd.: West Sussex, UK, 2007; 396.
6. Gilbert, M. K.; Luttrell, R. D.; Stout, D.; Vogt, F. Introducing chemometrics to the analytical curriculum: Combining theory and lab experience. *J. Chem. Educ.* **2008**, *85*(1), 135–137.
7. Wanke, R.; Stauffer, J. An advanced undergraduate chemistry laboratory experiment exploring nir spectroscopy and chemometrics. *J. Chem. Educ.* **2007**, *84*(7), 1171–1173.
8. Wang, L.; Mizaikoff, B.; Kranz, C. Quantification of sugar mixtures with near-infrared raman specroscopy and multivariate data analysis. *J. Chem. Educ.* **2009**, *86*(11), 1322–1325.
9. Rodriguez-Rodriguez, C.; Amigo, J. M.; Coello, J.; Maspoch, S. An introduction to multivariate curve resolution-alternating least squares: Spectrophotometric study of the acid-base equilibria of 8-hydroxyquinoline-5-sulfonic acid. *J. Chem. Educ.* **2007**, *84*(7), 1190–1192.
10. Sârbu, C.; Horovitz, O. Characterization and classification of lanthanides by multivariate-analysis methods. *J. Chem. Educ.* **2005**, *82*(3), 473–483.
11. Ribone, M. E.; Pagani, A. P.; Olivieri, A. C.; Goicoechea, H. C. Determination of the active principle in a syrup by spectrophotometry and principal component regression analysis. *J. Chem. Educ.* **2000**, *77*(10), 1330–1333.
12. Pojic, M. M.; Mastilovic, J. S.; Pestoric, M. V.; Dakovic, S. M. A comparative study of two analytical methods for fat content determination in brewer's grits. *J. Am. Soc. Brew. Chem.* **2009**, *67*(3), 166–169.
13. Alcala, M.; Ropero, J.; Vazquez, R.; Romanach, R. J. Deconvolution of chemical and physical information from intact tablets NIR spectra: Two- and three-way multivariate calibration strategies for drug quantitation. *J. Pharm. Sci.* **2009**, *98*(8), 2747–2758.
14. Gilpin, R. K.; Gilpin, C. S. Pharmaceuticals and related drugs. *Anal. Chem.* **2009**, *81*(12), 4679–4694.
15. Eaton, J. W.; Bateman, D.; Hauberg, S. *Gnu Octave Manual*; Network Theory Ltd.: U.K., 2008.
16. Haaland, D. M.; Easterling, R. G. Improved Sensitivity of Infrared Spectroscopy by the Application of Least Squares Methods. *Appl. Spectrosc.* **1980**, *34*(10), 539–548.
17. Thomas, E. V.; Haaland, D. M. Comparison of multivariate calibration methods for quantitative spectral analysis. *Anal. Chem.* **1990**, *62*(10), 1091–1099.
18. Anderson, E.; Bai, Z.; Bischof, C.; Blackford, S.; Demmel, J.; Dongarra, J.; Du Croz, J.; Greenbaum, A.; Hammarling, S.; McKenney, A.; Sorensen, D. *LAPACK User's Guide*, 3rd ed.; Society for Industrial and Applied Mathematics: Philadelphia, PA, USA, 1999.
19. Harris, D. C. *Quantitative Chemical Analysis*, 7th ed.; W. H. Freeman and Company: New York, 2007; 53–77.