



XVIIth World Congress of the International Commission of Agricultural and Biosystems Engineering (CIGR)

Hosted by the Canadian Society for Bioengineering (CSBE/SCGAB)
Québec City, Canada June 13-17, 2010



QUANTITATIVE ANALYSES ASPECT OF VIBRATIONAL SPECTROSCOPY AND ITS APPLICATIONS IN AGRICULTURE

WENBO WANG¹, JITENDRA PALIWAL²

¹ E2-376, EITC, University of Manitoba, R3T 5V6 umwangw2@cc.umanitoba.ca

² E2-376, EITC, University of Manitoba, R3T 5V6 J_Paliwal@umanitoba.ca

CSBE101162 – Presented at Section VII: Information Systems

ABSTRACT The three main branches of vibrational spectroscopy, i.e., mid-infrared, near-infrared, and Raman spectroscopy, are widely accepted techniques for qualitative analyses in the agriculture and food sectors. Mid-infrared and Raman spectroscopy probe the same ‘fingerprinting’ area for most organic as well inorganic matter. Such advantages make them ideal classification tools, with little need for sample preparation in routine analysis. Utilization of these vibrational spectroscopy techniques for quantitative analysis is more complicated and requires data handling to render meaningful interpretation of the spectra. A brief introduction to the theoretical background of infrared absorption and Raman inelastic scattering processes is presented. This article presents a detailed discussion of the fundamental principles behind quantitative analysis using absorption and emission spectroscopy techniques. Different measurement modes and their relationship to the acquisition and transformation of spectral data are explained with practical applications related to the quantitative theory. This paper also talks about the rationale for pre-treating the data. A review of the available procedures to pre-process different kinds of spectra and to build calibration concludes the paper. Linear calibration methods, e.g., classic linear regression, principal component regression, partial least squares regression, support vector machines, and nonlinear methods such as artificial neural network, are briefly reviewed.

Keywords: Vibrational spectroscopy, Raman, near-infrared, infrared, quantitative analysis, partial least squares regression, principal component regression.

INTRODUCTION Quality control and screening is an important part in food processing. Producers and processors in today’s food industry have plenty of choices when selecting the proper tools to meet their quality control needs. Vibrational spectroscopy, which probes energy transitions due to molecular vibrations, is a tool too valuable to skip. Compared to most lab methods, e.g., conventional wet chemistry and liquid chromatography, the advantage of vibrational spectroscopy lies in its speed and convenience. Little sample preparation is required and test on samples could be conducted in a non-destructive way. Types of samples classically measured by vibrational spectroscopy include gases, liquids, bulk and powdered solids. Nowadays, samples like interfacial species, microsamples, and trace analytes can also be analyzed on

modern spectrometers. As routine analytical methods, different branches of vibrational spectroscopy have achieved amazingly short measurement times that can vary from minutes to fractions of seconds (Griffiths and de Haseth 2007). It is, therefore, one of the most preferred tools for routine analysis of the chemical compositions of foods and agricultural products.

The Molecular vibrations give rise to absorption bands in the infrared region of electromagnetic spectrum. With incident photons of the correct frequency, the vibrational energy of molecules could be promoted to the next allowed energy level. These energy transitions are studied by popular absorption spectroscopy techniques such as mid-infrared (MIR) and near-infrared (NIR) spectroscopy. Besides these absorption techniques, Raman spectroscopy, which is based on inelastic scattering effect, has also proven to be a rather useful technique in food studies. Raman and infrared (IR) techniques are complementary techniques because their selection rules for molecular vibration are different. However, Raman shift and MIR absorption share the same energy range. One fundamental difference between absorption and Raman scattering is probability, with Raman effect being far less likely to happen, which makes Raman a weak phenomena (McCreery 2000). This lack of sensitivity can be remediated with special techniques such as resonance Raman or surface-enhanced Raman scattering (SERS).

Both the frequency and intensity of molecular vibrations induced by external radiation are subject to the chemistry and environment around the individual atoms. Thus, vibrational spectroscopy can be used to study chemical compositions of various food materials and products (Li-Chan, 1996). There exist a number of organic compounds and functional groups that are active to Raman and IR techniques. Each functional group could be identified by their unique pattern of either absorption or Raman scattering. The intensity of these spectral features can be used to calculate the relative concentration in the sample (Wetzel and LeVine 1999). Therefore, both qualitative and quantitative information of the target sample are obtainable using vibrational spectroscopy. The spectroscopic methods hold their distinct advantages over chemical analysis as a routine characterization tool. One significant advantage is that vibrational spectroscopy techniques require little to no sample preparation.

This paper describes the fundamental relationships in quantitative analysis using vibrational spectroscopy. Practical aspects such as sampling modes, data preprocessing methods, and calibration procedures are also reviewed.

BASIS FOR QUANTITATIVE ANALYSIS The theoretical explanations for the three branches of vibrational spectroscopy are different. However, the underlying principles for quantitative analysis are essentially the same. In absorption techniques such as NIR and IR, the measured spectral quantity is absorbance at individual wavelength variables. The values of the absorbance are correlated with certain properties of the sample, e.g., particle sizes and/or the concentration of the analyte. Particularly, for IR spectroscopy, the spectra offer unique absorption features at discrete wavelengths, which are called ‘fingerprinting’ areas for certain chemical species. This could be advantageous when using IR technique for qualitative analysis, where individual bands could be unambiguously assigned to each component in a mixed sample. To conduct quantitative analysis using IR spectroscopy, the change in the magnitude of the spectral signal could be directly correlated to the

chemical properties of sample under study. The absorption bands in NIR spectra are highly overlapping and often weak. Moreover, most NIR spectra of solid samples are conveniently acquired under diffuse reflection. Often full spectra are used with indirect calibration algorithms, e.g., partial least squares regression (PLSR), for dealing with NIR spectral data. Raman spectra offer good specificity for qualitative analysis because of the many well-resolved spectral features they have. Therefore, it is more common to use Raman spectroscopy for determination of molecular structures and chemical species discrimination rather than quantitative analysis. It is, however, worth mentioning that the use of Raman spectroscopy for quantitative analysis has been around for over 50 years (Pelletier 2003). Considering its great advantage in flexible sampling for many practical applications, Raman spectroscopy is a rather useful tool as an online analyzer especially under harsh environments.

Fundamental relationships The quantitative relation between the measured spectral signal, either through absorption or inelastic scattering, and the concentration of the analyte are defined here. In practice, most of the instruments do not take measurement under absorption mode directly but under the transmission mode instead. Transmission mode commonly applies to measurement of transparent or semi-transparent samples. For the ease of sampling of solid samples, diffuse reflectance mode is often applicable for all three vibrational spectroscopy techniques.

Laws of absorption The physical law that governs the absorption of electromagnetic radiation is usually called the Beer-Lambert law or the Beer-Lambert-Bouguer law. It consists of the following Beer's law and Lambert's (Bouguer's) law as explained in Fig.1.

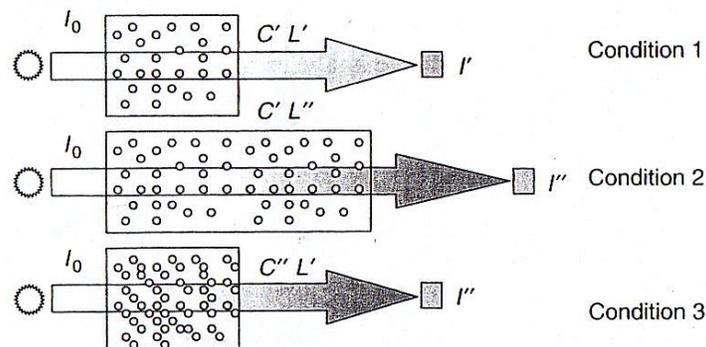


Figure 1. The relationships of absorption, concentration and optical pathlength. I_0 denotes the intensity of the incident beam; I' and I'' represents the transmitted energy; C' and C'' denotes the concentration of analytes; L' and L'' are the optical pathlengths of sample cell. Source: Coates (2002)

In condition 1 shown in Fig. 1, an incident beam with intensity of I_0 is projected perpendicularly onto the receiving interface of a sample cell. The concentration of analyte C' is defined as the number of molecules per unit volume. Passing through the sample material with an optical pathlength of L' , the beam of incident photons interacts with analyte molecules and are absorbed. On the exiting side of the sample cell, the intensity is attenuated to I' . If the optical pathlength is altered to double the amount of L' , the situation depicted in condition 2 applies. There, a new attenuated intensity value I'' is recorded. In another way, if the optical pathlength L' is kept constant, and the number of

molecules per unit volume is doubled, condition 3 occurs. It is noted here that the recorded attenuated intensity in condition 3 is the same as that derived under condition 2. Condition 1 and 2 describe the original Lambert's/Bouguer's law. The relation is expressed as follows:

$$A = k' \times L$$

where A is the measured absorbance, k' is a constant, and L is the optical pathlength. Basically, the Lambert's/Bouguer's law states that the absorbance changes proportionally with the optical pathlength provided that everything else remains the same.

If the thickness of the sample is set to a determined value and the concentration of analyte changes, the Beer's law describes the relation between absorbance and concentration. This is illustrated by condition 1 and condition 3 depicted in Fig. 1. The following equation states the Beer's law.

$$A = k'' \times C$$

where A is the measured absorbance, k'' is a constant for a given optical pathlength, and C is the concentration, i.e., molecules per unit volume. The Beer's law describes the proportional changes between absorbance and the concentration of analyte.

More commonly, people perceive the laws of absorption in the following combined form, which is called the Beer-Lambert-Bouguer law.

$$A = K \times C \times L$$

where A is the measured absorbance, K is a combined constant and is called molar absorptivity. Constant K is unique to the analyte species that absorbs radiation.

In many practical applications, the optical pathlength in the measurement is often kept constant. Therefore, Beer's law is employed instead of the combined law of absorption. It is also noted there are some assumptions for the law of absorption to hold. One assumption is that the electromagnetic radiation is monochromatic light. The other assumption is that the sample system is homogeneous. Deviation from these basic assumptions and relationships causes non-conformance to the ideal behaviours. However, such deviation does not necessarily affect the quantitative accuracy. The causes of discrepancies could be due to many parameters such as instrument parameters and the actual condition of samples.

Laws of Raman scattering In the treatise by Long (1977), a rather elaborate equation is given for the intensity of Raman scattered signal. In the book written by McCreery (2000), a classical and succinct form of the equation is given as follows:

$$I_R = I_0 \sigma_j D dz$$

where I_R is the intensity of measured Raman signal, I_0 is the intensity of excitation laser, σ_j is the Raman cross section, D is the number density of scatters (molecules per cm^3), and dz is the pathlength of excitation laser travelled in sample (cm). A first look at the equation shows that the relation is analogous to that described in Lambert-Beer-Bouguer's law in absorption spectroscopy. However, the pathlength term dz is kept constant for the sampling geometry.

Measurement modes The fundamentals for quantitative analysis with absorption and Raman scattering are explained in the previous section. However, the measurement of the variables in these equations is not as apparent as it seems. In absorption spectroscopy, transmittance is often taken instead of absorbance. The calculation of transmittance is based on measuring the amount of energy transmitted through samples and comparing it to the transmission of light without sample. Therefore, transmittance is a dimensionless ratio of transmitted energy compared to the energy in the absence of sample. The percentage scale of transmittance varies from 0 to 100. A zero transmittance represents total absorption while a value of 1 denotes no absorption. It is also noted that transmittance is not linearly proportional to absorbance, and hence not proportional to concentration of analyte. The relation between transmittance and absorbance is defined as follows:

$$T = \frac{I}{I_0}$$

$$A = -\log_{10} T = \log_{10} \frac{1}{T}$$

where T is transmittance, I is intensity of light transmitted through sample, and I_0 is transmitted energy without sample.

There are several possible ways to obtain spectral data. In Fig.2, some of the most common measurement modes are illustrated.

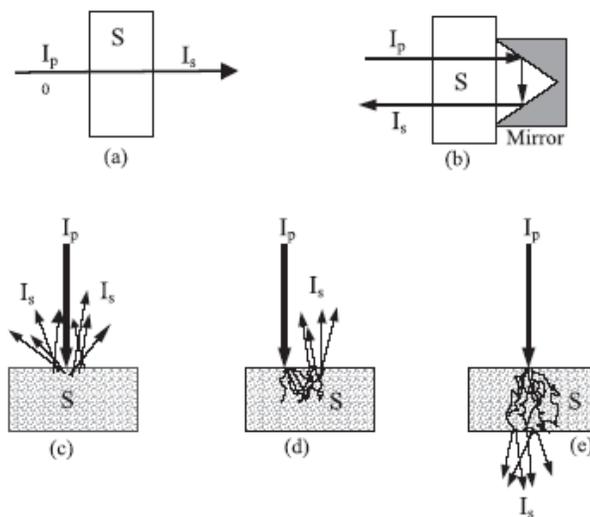


Figure 2. Modes of measurement a) transmittance; b) transflectance; c) diffuse reflectance; d) interactance; and e) diffuse transmittance through scattering medium. Source: Pasquini (2003)

As seen in the Figure, the first two modes, i.e, transmittance and transflectance, are the most appropriate methods to conduct transmission measurement. In IR spectroscopy, there is a sampling technique called attenuated total reflection (ATR). It is essentially a transmission measurement, though the spectral response is a function of incident angle and wavelength. Diffuse reflectance, interactance, and diffuse transmittance are most

appropriate sampling techniques for NIR measurement of solid samples. The reflectance data recorded in this way needs to be transformed. Methods to do this include the conversion to apparent absorbance, i.e., $\log_{10}(1/R)$, where R represents reflectance, and the Kubelka-Munk theory (Osborne 1993). Raman spectroscopy is a scattering technique and therefore reflectance measurement is the most appropriate.

DATA PROCESSING As discussed in previous sections, vibrational spectra of most compounds are fairly complex. Especially for NIR spectra, the absorption features are unintelligible and hard to interpret under most circumstances. Therefore, sometimes it may be necessary to preprocess the spectral data before applying calibration procedures. The purpose of this kind of data handling is to minimize the impact of irrelevant information embedded in the spectra. The noise signal, i.e., the irrelevant information, might be caused by many internal and external parameters such as light scattering process, system shift, detector spectral response, stray light, etc. Algorithms are proposed and developed to linearize spectra, remove baseline, and smooth data. Rinnan et al. (2009) provided an updated overview of the prevailing preprocessing techniques for handling NIR data. The methods mentioned in the document are universal and could be applied to pre-treat other types of vibrational spectra as needed. To make the Raman spectra suitable for quantitative analysis, Pelletier (2003) enlisted some of the useful preprocessing techniques. The collection of pre-processing methods includes common methods like Fourier transform, digital filtering, derivatives, co-adding, and curve fitting. It also contains methods specially designed for treating Raman data, e.g., spectra segmentation, bandpass filtering, etc. Although it appears appropriate to condition spectral data before quantitative analysis, caution must be exercised. Only the well-understood correction methods with minimal impact on quantitative aspect of spectra are suggested. It is also important to note that the same preprocessing method should be applied throughout the entire analysis process. That is to say, pre-treatment methods should be equally applied to spectral data used in calibration and spectra of samples in the prediction phase. For all vibrational spectroscopy, the spectral intensity could refer to the peak intensity, differential peak intensity, or integrated intensity area. More frequently, integrated intensity area is used in quantitative Raman analysis.

CALIBRATION Currently, the principal methods applied in multivariate calibration are multiple linear regression (MLR), principal components regression (PCR), and Partial least squares regression (PLSR) (Næs and Martens 1984). Martens and Næs (1984) explain some key concepts in multivariate calibration, e.g., direct vs. indirect calibration and controlled vs. natural calibration. Geladi and Kowalski (1986) provide a tutorial on PLSR and compared MLR, PCR, and PLSR. In practice, spectra for components in a mixture invariably overlap or are altered from spectral features of pure components, whether it is IR, Raman, or NIR. However, the principle of linear superposition works for Raman spectroscopy because Raman cross-sections are small. Therefore, methods like classical least squares (CLS), or sometimes called the K-matrix method, could be appropriate when spectra for the pure components are known. For NIRS applications, PCR and PLSR solve the collinearity problem in MLR. Compared to the two-step PCR, i.e., PCA and regression, PLSR also considers the inner relation between sample properties, e.g. compound concentrations, and spectral data during matrix decomposition. As a physicochemical method, measurement of NIR spectra is subject to deviations from Lambert-Beer-Bouguer law because of physical effects such as sample particle size, homogeneity, and temperature. These effects cause nonlinearity in collected data

(Borggaard 2001). Fuller et al. (1988) provided a comparative overview of various methods for multi-component quantitative analysis of IR data, i.e., K-matrix, multivariate least-squares, principal component regression (PCR), and partial least-squares (PLS). The advantages and disadvantages of each method were highlighted. Other than pre-treating data to linearize spectra response, nonlinear modeling methods such as ANNs can be used for classification and calibration in vibrational spectroscopy. The commonly used ANN architectures are back-propagation (BP) network, radial basis function (RBF) network, and Kohonen self-organizing map network (He et al. 2007; Dou et al. 2006; Massart et al. 1988). Support vector machine (SVM) is a relatively new classification technique that has only recently been introduced to the chemometrics arena to solve both classification and calibration problems (Xu et al. 2006). The advantages of SVM methods are their optimal generalization performance in ill-posed situations with limited amount of training data. Least-squares support vector machines (LS-SVM) has been recently introduced as a multivariate calibration technique (Thissen et al. 2004). However, for linear problems that are well-understood in analytical chemistry, SVMs are unnecessarily complicated and conventional approaches, such as LDA, PCA, and PLSR, are more effective (Xu et al. 2006).

CONCLUSION Vibrational spectroscopy, the study of interaction between electromagnetic radiation and molecules, is an important tool for quality inspection in the agriculture and food industries. In this paper, the fundamentals for three branches of vibrational spectroscopy are described. The central topic of this treatise is the use of vibrational spectroscopy for quantitative analysis. Therefore, effort is made to elaborate on the basic relationships between spectral signal and concentration of the analyte. Measurement modes employed for acquiring different spectral data are explained. Preprocessing and calibration methods for quantitative analysis are reviewed.

Acknowledgements. We would like to acknowledge the financial support provided by the Canada Graduate Scholarship funded by Natural Sciences and Engineering Research Council of Canada (NSERC) and Manitoba graduate scholarship through the Manitoba provincial government.

REFERENCES

- Borggaard, C. 2001. Neural Networks in Near-Infrared Spectroscopy. In *Near-infrared Technology in the Agricultural and Food Industries*, ed. P.C. Williams, K.H. Norris, 101-108, St. Paul, MN: AACC.
- Coates, J. 2002. Classical methods of quantitative analysis. In *Handbook of Vibrational Spectroscopy*, ed. J.M. Chalmers and P.R. Griffiths, 2241-2255, Chichester, West Sussex, England: Wiley.
- Dou, Y., H. Mi, L. Zhao, Y. Ren and Y. Ren. 2006. Radial basis function neural networks in non-destructive determination of compound aspirin tablets on NIR spectroscopy. *Spectrochim. Acta Part A*. 65(1): 79-83.
- Fuller, M.P., G.L. Ritter and C.S. Draper. 1988. Partial Least-Squares Quantitative Analysis of Infrared Spectroscopic Data. Part I: Algorithm Implementation. *Appl. Spectrosc.* 42(2): 199-369.
- Geladi, P. and B. Kowalski. 1986. Partial least-squares regression: A tutorial. *Analytica Chimica Acta* 185: 1-17.
- Griffiths, P.R. and J.A. de Haseth. 2007. *Fourier Transform Infrared Spectroscopy*. New

- York, NY: John Wiley and Sons.
- He, Y., X. Li and X. Deng. 2007. Discrimination of varieties of tea using near infrared spectroscopy by principal component analysis and BP model. *J. Food Eng.* 79(4): 1238-1242.
- Li-Chan, E.C.Y. 1996. The applications of Raman spectroscopy in food science. *Trends in Food Science and Technology* 7: 361-370.
- Long, D.A. 1977. *Raman spectroscopy*. New York: McGraw-Hill.
- Martens, H. and T. Næs. 1984. Multivariate calibration. I. Concepts and distinctions. *Trends Anal. Chem.* 3(8): 204-210.
- Massart, D.L., B.G.M. Vandeginste, S.N. Deming, Y. Michotte and L. Kaufman. 1988. *Chemometrics: A Textbook*. Amsterdam, Netherland: Elsevier.
- McCreery, R. L. 2000. *Raman Spectroscopy for Chemical Analysis*. New York, NY: John Wiley and Sons.
- Næs, T. and H. Martens. 1984. Multivariate calibration. II. Chemometric methods. *Trends Anal. Chem.* 3(10): 266-271.
- Osborne, B.G., T. Fearn and P.H. Hindle. 1993. *Practical NIR Spectroscopy: with Applications in Food and Beverage Analysis*, 2nd edition. Essex, England: Longman Scientific & Technical.
- Pasquin, C. 2003. Near Infrared Spectroscopy: Fundamentals, Practical Aspects and Analytical Applications. *Journal of Brazilian Chemistry Society* 14(2): 198-219.
- Pelletier, M.J. 2003. Quantitative analysis using Raman spectrometry. *Applied Spectroscopy* 57(1): 20A-39A.
- Rinnan, Å., F. van den Berg and S.B. Engelsen. 2009. Review of the most common pre-processing techniques for near-infrared spectra. *Trends in Analytical Chemistry* 28(10): 1201-1222.
- Thissen, U., B. Üstün, W. J. Melssen and L.M.C. Buydens. 2004. Multivariate Calibration with Least-Squares Support Vector Machines. *Anal. Chem.* 76(11): 3099-3105.
- Wetzel, D.L. and S.M. LeVine. 1999. Microspectroscopy — imaging molecular chemistry with infrared microscopy. *Science* 285: 1224–1225.
- Coombs, T. R., and F. C. Watson. 1997. *Computational Fluid Dynamics*. 3rd ed. Wageningen, The Netherlands: Elsevier Science.
- NSC. 2001. *Injury Facts Online*. Itasca, Ill.: National Safety Council. Available at: www.nsc.org. Accessed 17 December 2001.
- Waladi, W., B. Partek, and J. Manoosh. 1999. Regulating ammonia concentration in swine housing: Part II. Application examples. *Trans. ASAE* 43(4): 540-547.
- Xu, Y., S. Zomer and R.G. Brereton. 2006. Support vector machines: a recent method for classification in chemometrics. *Crit. Rev. Anal. Chem.* 36: 177-188.