Mathematical Methods to Analyze Spectroscopic Data − New Applications

5 authors, including:

Gustavo Nicolodelli
Federal University of Santa Catarina
80 PUBLICATIONS   **1,164** CITATIONS

Sebastião Pratavieira
University of São Paulo
135 PUBLICATIONS   **1,315** CITATIONS

V. S. Bagnato
University of São Paulo
809 PUBLICATIONS   **13,828** CITATIONS

Some of the authors of this publication are also working on these related projects:

Project   Spectroscopic techniques applied in agricultural samples View project

Project   Laser-induced breakdown spectroscopy applied to the nutritional study of soybean plants, healthy and Aphelenchoids sp. infected samples View project

Mathematical Methods to Analyze Spectroscopic Data – New Applications

E.S. Estracanholli, G. Nicolodelli, S. Pratavieira, C. Kurachi and V.S. Bagnato

1. Introduction

Absorption and fluorescence spectroscopies in the visible and/or infrared spectrum are good options when a fast and objective analysis is required. Spectroscopy is based on light-matter interactions. This interaction occurs in different ways, and each molecule or an ensemble of molecules will show a distinctive response. The vibrational spectroscopy provides a fingerprint of the vibrational levels of a molecule usually at mid-infrared (MIR) radiation ($400\text{-}4000\,cm^{-1}$). The optical spectroscopy uses the ultraviolet-visible (UV-VIS) region (200-1000nm) of the electromagnetic spectrum and interrogates the electronic levels of a molecule. The instrumentation used to generate and detect this radiation is less complex and cheaper compared to other spectroscopy techniques, such as nuclear magnetic resonance, X-rays, etc. An absorption spectrum is obtained by irradiating a sample and measuring the light which is transformed into other forms of energy, e.g. molecular vibration (heat). A fluorescence spectrum is obtained only from fluorescent molecules, those that absorb and then emit radiation, acquiring the intensity of light emitted as a function of the wavelength. These spectra are characteristic for each molecule, because each one has different electronic levels and vibrational modes. These levels and modes are also influenced by the solvent of the molecule.

Vibrational spectroscopy applied to the mid portion of the infrared spectrum provides the basis to develop several of the most powerful methods of qualitative and quantitative chemical analysis. Some of the advantages to use this technique are: the information is collected on a molecular level, almost any chemical group has IR bands, it is very environment-sensitive.[1-3]

Optical fluorescence spectroscopy is highly sensitive and can provide different information about the molecules and the molecular processes such as the molecular interaction with the environment, the molecular bonding and concentration.[3, 4]

INTECH
open science | open minds

Spectroscopic techniques in this range of the electromagnetic spectrum have shown applications in different areas, from analytical chemistry to the diagnosis of some types of cancer, detection of citrus diseases and of dental caries, with a high sensitivity and good specificity rates. This is possible because the analyzed systems are composed of different types and concentrations of molecules. Thus, the spectrum of samples obtained under different conditions will also be different. It is therefore possible to identify and also quantify different compounds. However, the spectral variation can be characterized and correlated only with difficulty. This is mainly due to the fact that other phenomena, such as scattering and/or absorption, happen with the emitted light. In some cases, there may be other molecules in the sample presenting absorption bands that overlap in the same spectral region of the compound of interest, this mainly happens for absorption spectroscopy. In other cases, as in the case of fluorescence spectroscopy, the excitation and emitted light can be absorbed by other molecules making the signal too weak to be detected. A solution to this problem may be statistical procedures applied where the spectral information is correlated with any parameter of interest. [3-5]

A new application of a statistical method to process multi-layer spectroscopy information will be presented in this chapter. A brief review of the mathematical methods to analyze these spectroscopy data will be shown here, followed by two distinct examples. The first example is UV-VIS fluorescence spectroscopy, applied to detect the postmortem interval (PMI) in an animal model. The spectroscopy and statistical methods of analysis presented can be extended to other samples, like food and beverage. Here, a MIR absorption spectroscopy of liquid samples will be presented to detect and quantify certain compounds during the production of beer. Another system to measure liquid samples, which consists of a sample holder, will also be presented. This system offers a cheaper technique with a better signal compared to techniques used to analyze liquid samples in the MIR region.

2. Pattern recognition in the complex spectral database – Example of fluorescence spectroscopy used in forensic medicine

One of the limitations of conventional methods to determine the Postmortem Interval (PMI) of an individual is the fact that the measurements cannot be performed in real time and *in situ*. Several factors, environmental and body conditions influence the tissue decomposition and the time evolution, resulting in a poor resolution. Considering this limitation to determine the PMI, a possible solution is a new more objective method based on a tissue characterization of the degradation phases through optical information using fluorescence spectroscopy. If proven sensitive enough, this method shows a main advantage over conventional methods: less inter-variance and quantitative tissue information. These characteristics are relevant because they are less influenced by individual skills.[6]

During the decomposition process a wide variety of organic materials are consumed by natural micro-organisms and other unknown compounds produced by them. Using an objective method based on the tissue characterization of the stages of degradation by means of optical information using ultraviolet-visible fluorescence spectroscopy, followed by a

statistical method based on PCA (Principal Component Analysis) made it possible to identify well features with time progression. The characteristic pattern of time evolution presented a high correlation coefficient, indicating that the chosen pattern presented a direct linear relationship with the time evolution. The results show the potential of fluorescence spectroscopy to determine the PMI, with at least a similar resolution compared to conventional methods.[6]

Another attractive feature of optical technologies is the fact that in situ information is achieved through a noninvasive and nondestructive interrogation with a fast response. Conventional laboratory techniques to determine PMI are time consuming and also require the cadaver removal from the location where it was found to a forensic lab facility. This operation already introduces additional changes to the analysis.

Fluorescence spectroscopy has been presented as a sensitive technique to biochemical and structural changes of tissues. The investigation of biological tissues is quite complex. Photons interact with biomolecules in several ways, and depending on the type of the interactions, they can be classified into three groups. The absorbers are the biomolecules that absorb photon energy. The fluorophores are biomolecules that absorb and emit fluorescent light. The scatterers are biomolecules that do not absorb the photons but change their direction. Several endogenous cromophores contribute and modify the final tissue spectrum. Distinct fluorophores emit light but the collected spectrum will be modified depending on the presence of absorbers and scatterers in the microenvironment on the path of the emitted photons and the probe interrogator. Taking into account all these light interactions that occur within the biological tissues, it is important to keep in mind that a tissue fluorescence spectrum is a result of the combination of all these processes occurring in the pathway between excitation and collection: excitation absorption and scattering, fluorescence emission, and fluorescence absorption and scattering.[5, 7]

Tissue changes begin to take place in the cadaver as soon as there is cessation of life. Optical characteristics change, and these changes may be detected using fluorescence spectroscopy. With the cessation of the metabolic reactions, tissue modifications are induced by several distinct factors, e.g. lack of oxygen and adenosine triphosphate, and intestinal microorganism proliferation. In this type of analyses, we first aimed to establish a proof of concept that fluorescence spectral variations for distinct PMIs are higher than the variance observed within each PMI. If the results are positive, a spectral time behavior can then be determined, i.e. fluorescence spectral changes identifying each PMI. This proposed method can be used to determine an unknown PMI based on a comparative analysis of a spectral database pattern. There is a potential correlation of the tissue fluorescence changes and the PMI, even though the same limitations concerning the time course variability of the cadaveric phenomena are also present, the optical spectra information can provide a more objective estimation.

Taking into account the resolution limitation to determine PMI in biological tissues, where the degradation process is non-homogenous and influenced by environment and cadaver intrinsic factors, the optical techniques may show a better PMI prediction when compared to current techniques.

The process using the principal component analysis is necessary to change the space analysis. For any type of spectroscopy, where space is determined by an analysis of light intensity in wavelengths of absorption or fluorescence, a change of base can be accomplished, where the variables become the variance of the dataset. We can explore this idea mathematically using a practical example. The set of spectra shown in figure 1 is a typical result of multiple samples.
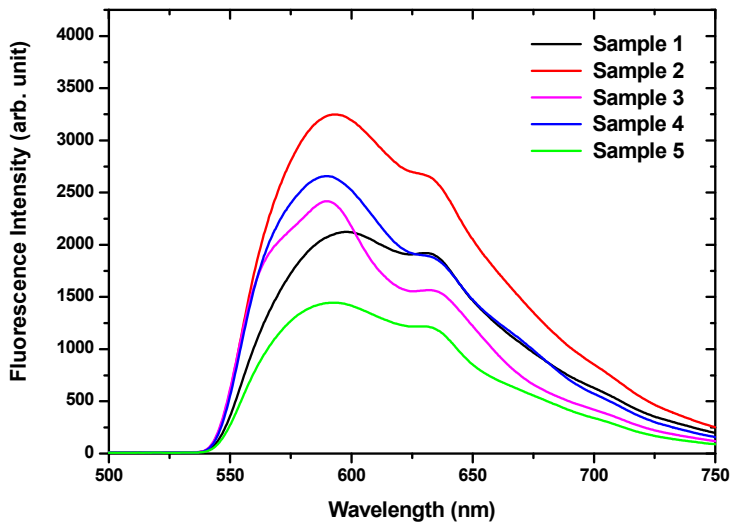
Figure 1. Fluorescence spectra of different samples.

In this case, we are dealing with intensities as a function of the wavelength, presented in a graphic form. These same curves can be represented in a matrix, as shown in table 1. In this form, each row corresponds to a single measurement, i.e., spectrum of a sample, and each column is the value of the wavelength considered, which makes up this spectrum. Thus, each array element is the intensity measured at a wavelength specific to a spectrum.

Wavelength (nm)	540	541	542	543	...	...	749	750
Intensity Sample1	27.5	30.6	33.9	37.7	...	...	199.9	198.4
Intensity Sample2	25.5	26.7	29.6	35.9	...	...	180.5	186.3
...	...	...	...	...	...	...	...	...
Intensity SampleN	24.1	25.3	27.9	30.2	...	...	176.8	180.6

Table 1. Fluorescence spectra in a matrix.

The next procedure to be performed after inserting the data set into this matrix representation is the centralization of the data around their average value. By fixing one column (wavelength) at a time we can calculate the average value for all lines (samples). Table 2 shows the mean values obtained.

Wavelength (nm)	540	541	542	543	...	...	749	750
Intensity Media	25.7	27.53	30.47	34.6	...	...	185.73	188.43

Table 2. Matrix of mean values for each wavelength.

In the next step, each of the intensity values of each sample should be subtracted from this average value in the respective wavelength. The results for our example are shown in table 3.

Wavelength (nm)	540	541	542	543	...	...	749	750
Intensity Sample1	1.8	3.07	3.43	3.1	...	...	14.17	9.97
Intensity Sample2	-0.2	-0.83	-0.87	1.3	...	...	-5.23	-2.13
...	...	...	...	...	...	...	...	...
Intensity SampleN	-1.6	-2.23	-2.57	-4.4	...	...	-8.93	-7.83

Table 3. Normalized fluorescence spectra.

It is important to note that this procedure resulted in a better match between the variables. The first values had higher intensities (approximately 8 times) than the wavelengths around 750nm in relation to values around 540 nm. If this normalization of the data had not been performed, the outcome would have had a greater influence for the longer wavelengths, as if they possessed some kind of "privilege", which would not be correct from the standpoint that all the measured variables are also important.

Since each element of the initial data array is represented by an element q_{ij}, we can consider this procedure performed using the equation 1:

$$X_{ij} = q_{ij} - \bar{q}_j \tag{1}$$

Where: x_{ij} is the element of our new data matrix; q_{ij} is the array element data corresponding to the i-th measurement variable j; $\bar{q}_j$ is the mean value of the variable j;

As the data were previously normalized, i.e. centered on their mean values, we proceed with the construction of the correlation matrix, where we obtain information about a dataset that indicates how the variables are correlated. This is possible by calculating the product of the transposed data matrix by itself. In mathematical terms, if x is our new array of standardized data x_{ij} composed of elements, then the correlation matrix R formed by these correlation coefficients is given by:

$$\mathbf{R} = \mathbf{X}^T \cdot \mathbf{X} \tag{2}$$

A matrix whose elements are given by:

$$r_{jj'} = \sum_{i=1}^{n} x_{ij} x_{ij'} = \sum_{i=1}^{n} \frac{(q_{ij} - \bar{q}_j).(q_{ij'} - \bar{q}_{j'})}{\sigma_j . \sigma_{j'}} \tag{3}$$

The r_{jj} value is a standardized covariance between -1 and 1. It should be noted that the matrix is Hermitian (symmetric in the case of real variables, which is our case). We can also confirm that the elements along the main diagonal of the correlation matrix (elements where j = j′) correspond to the variance of the variable q_j. As noted earlier, the correlation matrix R is Hermitian and therefore its eigenvalues are real and positive and its eigenvectors are orthogonal. For this procedure, it is important to note that the values of the wavelengths themselves are not taken into account in the mathematical calculation. The data selected for the next step are the rows and columns highlighted in table 3.

After calculating all the elements of the correlation matrix, the diagonalization is necessary. The diagonalization process provides two sets of data. The first are the eigenvectors: vectors which constitute a new base having the direction and sense in which the initial data set has more tendencies to vary, i.e. the maximization of the variance. The second sets of data are the eigenvalues, which provide the weight information, i.e. the relevance of each of the directions of the eigenvectors.

The eigenvalues are represented by the matrix K and the eigenvectors by the matrix V:

$$K = \begin{bmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ ... & ... & ... & 0 \\ 0 & 0 & ... & \lambda_n \end{bmatrix} \tag{4}$$

As the diagonalized matrix was a correlation matrix, we assume that this new base formed by the set of eigenvectors of each R represents a percentage of the total variance; and the information contained in each eigenvector are unique and exclusive, since they are mutually orthogonal. Each λ_i in the K matrix represent the weight of the specific eigenvector. Through these eigenvalues we can determine which of the principal components explains the greatest amount of data. For the simple relationship between the value of each eigenvalue divided by the sum of all eigenvalues, i.e. $\lambda_i / \sum \lambda_i$, we can determine the weight (or representation) of each eigenvector.

Once we determined the basis that maximizes the variance of the data set, which should be done by "projecting" the initial data matrix in this new basis through the product between the eigenvectors and the matrix of normalized data, we obtain:

$$\mathbf{S=V \cdot Q} \tag{5}$$

This data set designed in the new base (matrix S) is known as *Score*. The transformation of the basic matrix data S into the matrix Q by the base which maximizes the variances is known as Karhunen-Loève transformation.

The matrix of scores representing the data in our new base is expressed in such a way that each column represents the projection of the initial data into one of the eigenvectors, or in other terms, in either direction variance. Each line of the S matrix still represents a measure, or spectrum, as shown in table 4.

	PC1	PC2	PC3	PC4	...	...	PC N-1	PCN
Sample1	0.6	0.4	0.7	1.1	...	...	2.0	1.7
Sample2	0.8	1.2	1.4	2.2	...	...	1.7	0.6
...	...	...	...	...	...	...	...	...
SampleN	0.2	2.1	0.1	0.3	...	...	0.9	0.5

Table 4. Presentation of the available data with the new basis

Now, instead of analyzing the data obtained on the basis of the variables that were defined as the value of the intensity at each wavelength, these are considered in the space of the variances of these values. This change of base allows a significant reduction of the information in which the data are analyzed.

Spectroscopy experiments measure intensity values in hundreds or thousands of wavelengths, providing up to hundreds or thousands of variables to be analyzed. Depending on the experiment, when calculating the principal components of this system we represent around 90% of the information system, i.e. the spectra obtained in only two components of this new base. The graph of two major principal components (PC1 versus PC2) provides a much better view to then analyze these data rather than the hundreds of dimensions we had obtained before. In other words, we now work with a significantly reduced number of variables, wavelengths, with no loss of information. On this new basis, each sample, which was previously represented graphically by a curve with hundreds of points, shall be represented by a single point only. This significant reduction and simplification makes it much easier to detect spectral patterns.

We will now go back to the example of determining the postmortem interval - the set of measures shown in figure 1. Each sample is a fluorescence spectrum from different postmortem intervals. We apply the above procedure and obtain results on this new basis. For this case, the first two principal components show a representation greater than 91%. These results are shown in figure 2, where each point represents a spectrum.

Comparing data of figure 2 to those of figure 1 a temporal evolution of standard measurements becomes obvious. Based on this analysis, each region of space PC1 x PC2 is characterized by a postmortem interval. So in terms of a practical application, if we have a spectrum obtained from an unknown postmortem interval, we just design it based on this new basis and thus allow matching of the region of space to this spectrum, represented by a new point. Thus, this procedure can be used to determine a postmortem interval using fluorescence spectroscopy for situations where this value is unknown.

The present methodology and results shown by Estracanholli et al.[6, 8] demonstrated the use of fluorescence tissue spectroscopy to determine PMI as a valuable tool in forensic medicine. Two approaches were employed to associate the spectral changes with the time evolution of tissue modification. First, direct spectral changes were computed using inter-spectra analysis, allowing establishing a pattern of the sample distribution with a time evolution. Second, the use of a statistical method based on PCA helped to identify features over time. In both cases, the characteristic time evolution pattern presented a high

correlation coefficient, indicating that the chosen pattern presented a direct linear relationship with time.
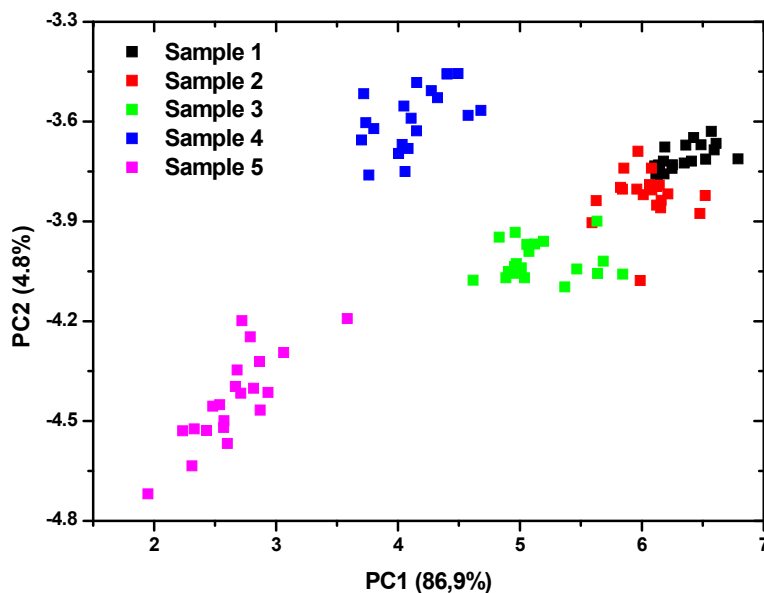
Figure 2. PC1 versus PC2 showing a more evident distinction between the samples.

However, other cases of application of spectroscopic techniques require a more robust processing. One such case is the other example cited above, where the goal is to quantify various compounds in a complex sample containing various interferences, and this most often occurs in the regions of overlapping absorption (or transmission and emission) of compounds of interest. For these cases, the application of artificial neural networks is a powerful solution.

3. Quantification of a composite with overlapping bands - Using vibrational spectroscopy in a beer sample

Beer brewing is a relatively long and complex biotechnological process, which can generate a range of products with distinct quality and organoleptic characteristics, all of which may be relevant to determine the type of product that should be made. Failures during important steps such as saccharification and fermentation can lead to major financial losses, i.e. to a loss of a whole batch of beer. Currently, analyses of the physic-chemical processes are carried out offline using traditional tests which do not provide any immediate response, e.g. HPLC (High Performance Liquid Chromatography). In the case of micro- breweries, which currently increase, some of these tests cannot be performed at all, due to the prohibitive cost of these tests. Therefore, many breweries do not have a possibility to identify errors during the production and to take corrective actions early-on. Today, problems are detected only at

a later stage, towards the end of the brewing process. Currently, most systems used in the breweries consume time and can potentially compromise the quality of a whole batch. A solution to this problem is a new method consisting of a system to monitor in real time (online) the saccharification and fermentation steps of the wort. The amounts of alpha-amylase and beta-amylase in the grain are correlated with the time required to convert all grain starch into sugars. [3, 9-11]

The brewing-process is based on traditional recipes, a defined period of time and temperature. The amount of different types of enzymes in the grain when the wort is produced is, however, not known, since this amount depends on many factors, e.g. storage conditions, temperature, humidity, transport. Due to these factors, the saccharification step could be stopped, which would mean that a significant amount of starch would remain in the wort, and therefore the procedure would result in poor wort. Or, it is also possible that all starch may have been converted to sugar and that the process continues longer than necessary. It is therefore critical to obtain data concerning the amount of sugar and alcohol in the wort fast. It is possible to get these data, using absorbance data in the mid infrared region (MIR) and analyzing these statistically using PCA and Artificial Neural Network (ANN) to determine the amount of sugars and alcohol in the wort during the saccharification and the fermentation procedure. These optical techniques provide huge advantages because they can be easily adapted to the industrial equipment, providing real-time responses with a high specificity and sensitivity. By applying these techniques, the procedure of saccharification and fermentation can be modified in each brewing step to increase the quality of the wort and eventually of the beer. This routine analysis during processing can also be used for other liquid samples.

A main feature of ANN is its ability to learn from examples, without having been specifically programmed in a certain way. In the case of spectroscopy, satisfying results can be achieved when ANN is used with supervised training algorithms. The external supervisor (researcher) provides information about the desired response for the input patterns, i.e. where there is an "a priori knowledge" of the problem. A neural network can be defined as applying non-linear vector spaces between input and output. This is done through layers of neurons and activation functions, where the input values are added according to weight and "bias" specific, producing a single output value [12-14]. A network "feedforward" is progressive or shows no recursion, if the input vector and a layer formed by the values precede the output layer, as shown in figure 3.

Formally, the activation function of the i-th neuron in the j-th layer is denoted by $F_{i,j}(\times)$; its output itself, j, can be calculated from the output of the previous layer itself, j-1, the weights $W_{i,k,j-1}$ (the index k indicates the neuron connected to the preceding layer) and bias $b_{i,j}$ according to the following formula

$$s_{i,j} = F_{i,j}\left(b_{i,j} + \sum_{k} w_{i,k,j-1}\, s_{k,j-1} \right) \tag{6}$$
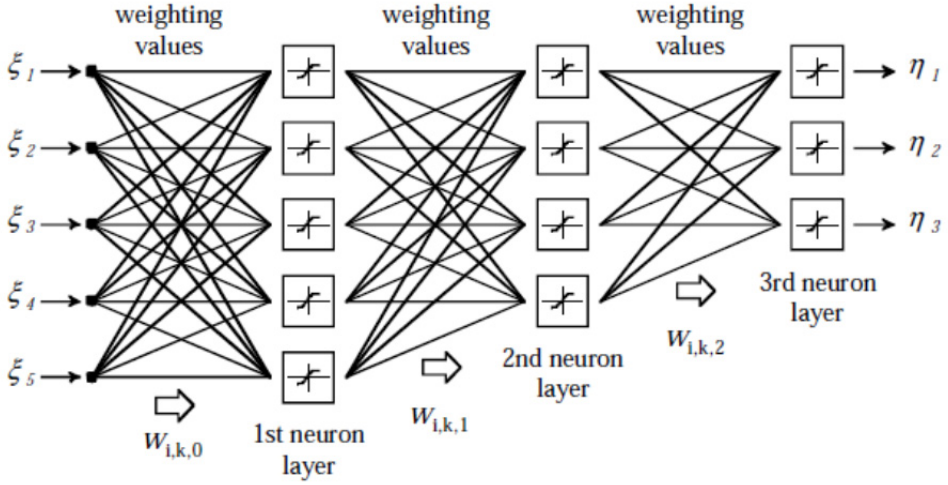
Figure 3. Schematic architecture of a neural network (perceptron multilayer).

The input and output values of the network being denoted by ξ_i and η_i respectively, the mapping can be determined due to a successive application of equation 6, which results for example in the following equation in the previous case:

$$\xi_i = F_{i,3}\left(b_{i,3} + \sum_{k=1}^{4} w_{i,k,2}\, F_{k,2}\left(b_{k,2} + \sum_{m=1}^{5} w_{k,m,1} F_{m,1}\left(b_{m,1} + \sum_{n=1}^{5} w_{m,n,0}\, \eta_n\right)\right)\right) \tag{7}$$

Since the choice of the activation function usually falls on the logistic sigmoid due to some of its mathematical properties (be class C ∞, for example), the above expression shows the relationship between ξ_i and η_i wich is defined by the weighing values and the bias. A very important characteristic of NN is its ability to learn, or the ability to reproduce the input-output pairs predetermined by properly adjusting the weights and the bias from training data and according to an adjustment rule. The method of a *"backpropagation"* rule is probably the best known training, and it is especially suited for progressive architectures. This rule is based on the successive application of the maximum slope algorithm determined from the first derivatives of the error between the desired outputs obtained by the parameters of the internal network. The backpropagation can be summarized in the following steps: (1) initialize the network parameters, $b_{i,j}$ and $w_{i,k,j}$ (2) select an entry ξ_i^p training data and form the pair (η_i^p, δ_i^p) , (3) calculate the error with a standard convenient Euclidean, e.g.

$$e = \sqrt{\Sigma_i(\delta_i^p - \eta_i^p)^2} \tag{8}$$

(4) Calculate the error derived from the above equation in relation to $b_{i,j}$ and $w_{i,k,j}$ (5) modify the parameters of the network according to the following rule and learning rate:

$$b_{i,j} \; \leftarrow \; b_{i,j} - \alpha \frac{\partial e}{\partial b_{i,j}} \quad and \quad w_{i,k,j} \; \leftarrow \; w_{i,k,j} - \alpha \frac{\partial e}{\partial w_{i,k,j}} \tag{9}$$

(6) Iterate steps (2) through (5) until a number of training cycles or stopping criteria has been achieved.[12, 13, 15, 16]

We can show in our case of beer analysis to which extent this processing technique is powerful. It has been applied widely in the interpretation of spectral data. In this case, an infrared absorption spectrum was obtained by Fourier Transform Infrared (FTIR) spectrometer [1, 2], the spectra is show in figure 4 and figure 5. In this case, the research objective was to provide a new method to determine the concentration of sugars and ethanol in beer wort during beer saccharification and fermentation in a short processing time. In our example, compounds of interest to be quantified can be separated into four main types of sugars present in the sample: glucose, maltose, maltotriose, dextrin (sugar chain length) and ethanol. It is important to note that the maltose binding is composed of two molecules of glucose, maltotriose three molecules of glucose sugar and that dextrins are composed of a large number of glucoses. Thus, the fundamental basis of these sugars is the same, the glucose, being differentiated only by the number of basic elements connected.

The absorption bands of these elements are expected to be so close that there is an overlap in the spectra, making the detection and quantification very complex. Figure 4 shows an example of absorption spectrum of a sample of ethanol, maltose 10%, and beer wort which contains some types of sugars. It is quite difficult to distinguish between the absorption spectra of the beer wort and the maltose, which contains certain types of sugars.

If we consider also the presence of ethanol (which has an absorption band in the same spectral region as the sugar) in the fermentation step, the procedure becomes even more complex. In figure 5 the extent of absorption during the fermentation step is shown, where the sample had initially all sugars without ethanol and ends up having only a part of dextrin (no fermentable sugar) and ethanol.

In this case, we first use the technique of principal component analysis in order to achieve a reduction of the number of variables to be analyzed. These spectra, which originally had about 1000 variables (wavelength where the absorbance is measured), can by these means represented by a few (two, in this case) variables, or principal components, with a high representation of information: 97.9%. The relationship between the two higher principal components is presented in figure 6 below.

Each spectrum of figure 5 is represented in figure 6 by a single point. In this new base of analysis, the wavelengths do not have any more significance, but the variance is important now. Each pair (PC1, PC2) represents a specific concentration of sugar and ethanol, which changes during the fermentation process. It is computationally feasible at this time, to apply an artificial neural network based on the values of the pairs for each of these points. For the first time, this experiment should be performed as previously described: as a case of a supervised NN e.g. a multilayer perception network). Therefore a method is required as the

gold standard to calibrate or to train our neural network. One of the most widely accepted methods is the technique of HPLC (High Performance Liquid Chromatography). Using this technique, we can accurately quantify all the types of sugars of interest and the ethanol. The compounds of interest were measured using the standard method and assembling the ANN. In the neural network input (η_i) using ordered pairs of principal components and the output (ξ_i), the results obtained using the HPLC techniques show the amounts of compounds of interest, in this case, the sugars and the ethanol.
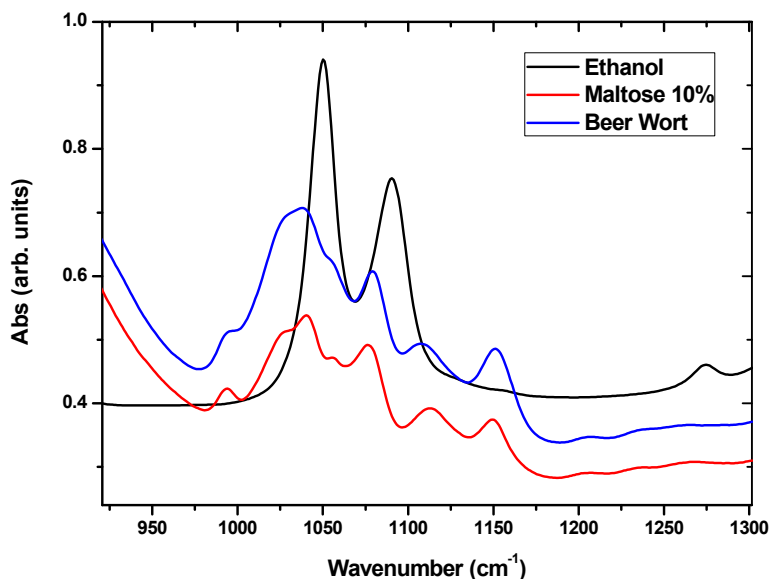
Figure 4. Absorption spectrum of a sample of ethanol, maltose 10%, and beer wort which contains certain types of sugars.

A certain part of the data (approximately 1/3), must be separated first in order to perform a further validation step. With 2/3 of the remaining data, the neural network is performed in the training stage following the equations and the structures described before, where the weight of each neural layer is adjusted in order to converge the network. The adjustment can be done as often as necessary, until the output (ξ_i) is as close to the true (real) value as required.

The training is complete, oncet the weight of the neurons has been adjusted, and the network has been converged with the desired error. The weight values should then be saved and stored before proceeding to the next step, which is the validation step: using the neural network to provide results of new spectra. With the data that were originally separated (1/3 data) and the values of the weights defined by the training stage, the neural network is run again. At this stage, note that the backpropagation system should not be executed. Simply use the matrix of the weights saved, and insert the data that have been separated for this validation step as inputs for the new network. Thus, the network will be performed only in the forward direction, supplying in a very short time of processing the output values ξ_i.

These output values are compared with the expected values using the HPLC technique, using a correlation curve between the two techniques. If these results are satisfactory, the process of mounting the system to quantify the compounds of interest is complete, and can be passed on for practical use.
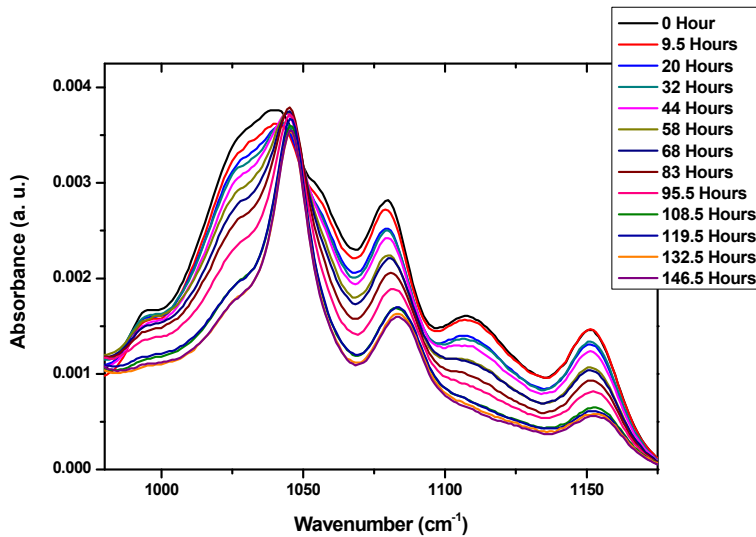

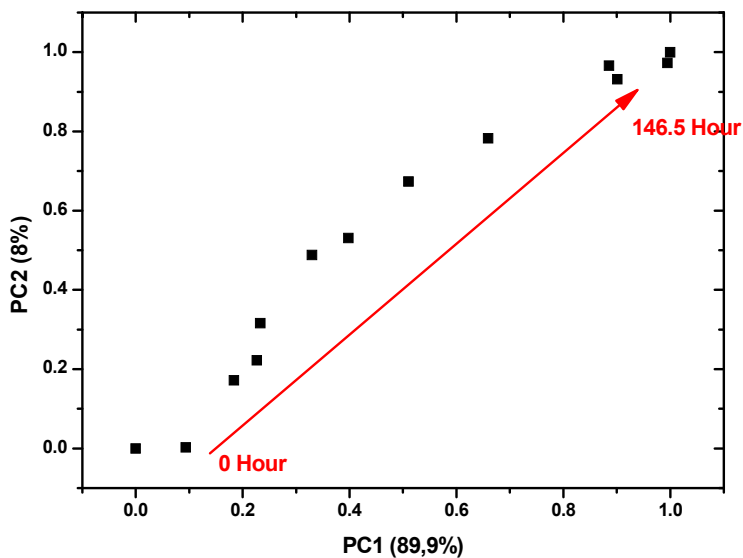**Figure 5.** Absorption during the fermentation process.

Figure 6. PC1 versus PC2 showing the time evolution of the fermentation process.

To be able to use spectroscopy with the neural processing requires using a standard method. We can simply use the PCA to reduce processing variables, entering the values of ordered pairs into the network, together with the weight values and collecting predetermined output results, in this case the amounts of sugars and ethanol. In the case of fermentation of the wort, using a number of principal components around three , a neural network comprising an input layer with 23 neurons and an output layer of 5 neurons , it is possible to quantify each type of sugar and ethanol with a quoted error of ± 0.2%. Here we exemplify our results showing the correlation between the value determined by the concentration of maltose using spectroscopy and HPLC technique (figure 7), where R^2 and the coefficient slope is 0.991 and 0.999 respectively. The results of a linear fit show a good agreement between the proposed new method and the standard procedure. This result allows the use of our technique in brewery, as it enables monitoring quality and making process control less time consuming.
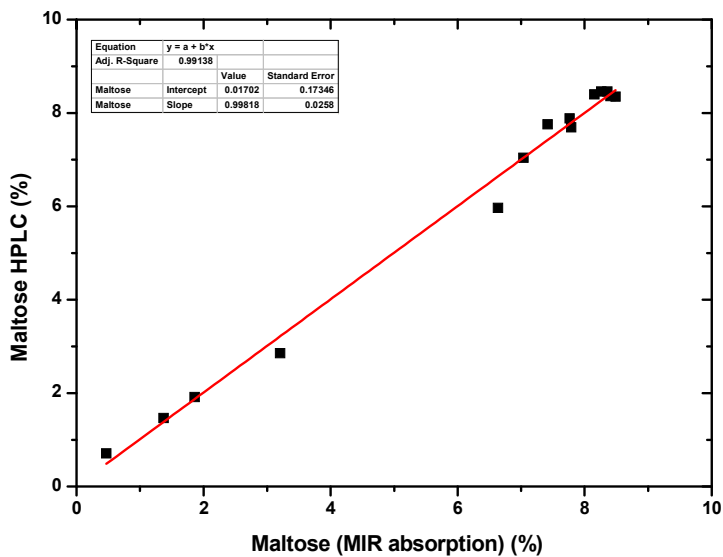
Figure 7. Correlation between the standard method (HPLC) and proposed procedure (MIR absorption).

4. Conclusion

In the analysis of spectroscopic data, not only the technique to obtain the values of different properties is important, but the correct mathematical processing of the data is actually the main issue to obtain the correct information. Especially the distinctions of multiple values which are correlated to a specific class of phenomena are the hide information that can be conveniently extracted. During our exposition in this chapter, we have concentrated in demonstrating how powerful the correct spectroscopy analysis can be when the first obtained data have been correctly arranged, allowing a mathematical procedure that treats the information as a whole instead of concentrations in individual values. Many techniques

are today available for such procedures, but especially the Principal Component Analysis is quite powerful to be applied when spectral information is not restricted to a single wavelength, but rather to a large portion of the spectra.

We have concentrated on a relevant case where the UV-VIS portion of fluorescence spectrum is obtained and applied to determine its correlation with the postmortem interval in an animal model. The fluorescence in this case is subject to many effects due to the biological tissue modification as a natural evolution once the living metabolic action has been interrupted. This is clearly the case where biochemical modification causes alteration of spectrum as a whole and the attempt to concentrate the observation on individual features may fail. With the application of PCA to collect data, rich information patterns made a high correlation between extracted information and the real postmortem time interval possible. The classification of patterns and congregations of collections of information create a distinction into groups of distinct PMI. Even though we have used the method for PMI determination, the method has been shown to be as well powerful in applications in the field of cancer diagnostic, fermentation processing in beverage production, quality control in industry, identification of plagues and other features of interest in agriculture. The level of application of the PCA technique can go beyond the identification of pattern and correlation with values and can also provide specific quantification of individual chemical components of the system which is investigated.

To demonstrate this feature, we consider as an example the sugar quantification during beer production. These cases represent a bigger challenge to innumerous systems in several areas. Using the PCA procedure associated with a Neural Network (NN) we can quantify the composites in a sample, obtaining results comparably quickly. Here, we used the example of beer analysis. Using the MIR absorption spectroscopy of liquid samples, without any type of pre-procedure, we detected and quantified specific compounds (glucose, maltose, maltotriose, dextrin and ethanol) during the production of beer. The NN were used to determine the amount of these types of sugar and alcohol in the wort during the saccharification and fermentation. In the correlation between the values determined by the concentration of maltose spectroscopy with the HPLC technique we find the $R2$ and coefficient slope to be 0.991 and 0.999 respectively. Finally, the presentation of this chapter is to show the real power of the conjugation of spectroscopy techniques with data analyses. The field is clearly growing in diversity and importance.

Author details

E.S.Estracanholli, G.Nicolodelli, S.Pratavieira, C.Kurachi and V.S. Bagnato
Institute of Physics of São Carlos, University of São Paulo, SP, Brazil

Acknowledgement

5. References

[1]   Socrates G. Infrared and raman characteristic: group frequencies: J. Wiley 2001.

[2]   Stuart BH. Infrared spectroscopy: Fundamentals and applications: J. Wiley 2004.

[3]   Bauman RP. Absorption spectroscopy: John Wiley 1963.

[4]   Lakowicz J. Principles of Fluorescence Spectroscopy. Springer 2006.

[5]   Ramanujam N. Fluorescence spectroscopy of neoplastic and non-neoplastic tissues. Neoplasia. 2000 Jan-Apr;2(1-2):89-117.

[6]   Estracanholli ES, Kurachi C, Bagnato VS. Tissue Fluorescence Spectroscopy in Death Time Estimation. *Forensic Pathology Reviews*: Springer Science 2011.

[7]   Pratavieira S, Andrade CT, Salvio AG, Bagnato VS, Kurachi C. Optical Imaging as Auxiliary Tool in Skin Cancer Diagnosis. *Skin Cancers - Risk Factors, Prevention and Therapy* 2011.

[8]   Estracanholli ES, Kurachi C, Vicente JR, de Menezes PFC, Silva OCE, Bagnato VS. Determination of post-mortem interval using in situ tissue optical fluorescence. Optics Express. 2009 May 11;17(10):8185-92.

[9]   Schieberle HDBWGP. Food chemistry: Springer 2009.

[10]   Woodcock T, Downey G, O'Donnell CP. Better quality food and beverages: the role of near infrared spectroscopy. Journal of near Infrared Spectroscopy. 2008;16(1):1-29.

[11]   Bamforth CW. Food, fermentation and Micro-organisms. : Blackwell Science 2005.

[12]   Kohonen T. Analysis of a Simple Self-Organizing Process. Biological Cybernetics. 1982;44(2):135-40.

[13]   Kohonen T. Self-Organized Formation of Topologically Correct Feature Maps. Biological Cybernetics. 1982;43(1):59-69.

[14]   Kandel ER, Schwartz JH, Jessel TM. Principles of neural science: McGraw-Hill 2000.

[15]   Lefebvre JCPNREWC. Neural and Adaptive Systems: Fundamentals through Simulations: John Wiley & Sons 1999.

[16]   Marks RDRRJ. Neural smithing: supervised learning in feedforward artificial neural networks. : MIT Press 1999.