

# Olive Oil Adulteration with Sunflower and Corn Oil Using Molecular Fluorescence Spectroscopy

Betül Öztürk, Aysun Ankan and Durmuş Özdemir

*İzmir Institute of Technology, Faculty of Science, Department of Chemistry, Gülbahçe, Urla, İzmir, Turkey*

## 50.1 INTRODUCTION

Olive oil is a valuable food product as compared with other vegetable oils due to its distinct taste, flavor and possible health benefits. The economic value of olive oil is generally much higher than other seed oils. As a result, the adulteration of olive oil with cheaper vegetable oils becomes a real concern. For this reason, the analysis of edible oils for possible adulterants is very important for food safety and protection of consumers. Based on the extraction method used, there are various types of olive oil on the market today. Extra virgin olive oil is obtained from the olive by purely mechanical means, and the lower grade oils are obtained by solvent extraction, heat treatment, esterification or refining. The composition of the oils is based on the fatty acids present and their locations on the glycerol backbone. This composition varies not only with the type of oil and extraction method but also with the geographical origin and meteorological effects during the growth and harvest of the olives (Tay et al., 2002). This variation can be used for oil authentication and the identification of adulteration. Various physical and chemical tests have been used to establish the authenticity of olive oil and to detect the level of adulterants in it (Aparicio et al., 1997; Dennis, 1998; Christopoulou et al., 2004). Studies related to olive oil adulteration were mostly carried out with chromatographic methods in recent years (Wenzla et al., 2002; Ghosh et al., 2005; Hajimahmoodi et al., 2005). However, while chromatographic methods offer high sensitivity and accuracy, they are also time-consuming and expensive. On the other hand, spectroscopic methods may offer faster and cheaper analysis alternatives (Papadopoulos et al., 2002; Schulz et al., 2003; Guimet et al., 2005).

Molecular fluorescence spectroscopy is a sensitive technique for differentiating various seed oils from olive oils as

they exhibit natural fluorescence (Kyriakidis and Skarkalis, 2000). However, when multicomponent oil mixtures need to be quantified, somewhat broad overlapping excitation and emission peaks make it necessary to use multivariate calibration methods. Nevertheless, there have been an increasing number of reports in recent years about the use of excitation-emission fluorescence (EEF) spectroscopy and synchronous fluorescence (SF) spectroscopy for the determination of olive oil adulteration with cheaper oils (Francesca et al., 2004a,b; Konstantina et al., 2006). In EEF, an excitation-emission matrix (EEM) which consists of emission spectra measured at different excitation wavelengths, is recorded for a sample. In SF mode, the excitation wavelength is increased with a constant wavelength increment while the emission spectrum is recorded in a predefined wavelength range, thus a constant wavelength interval is maintained between excitation and emission wavelengths ( $\Delta\lambda$ ) as the scan progresses. As a result, the SF method produces a two-dimensional fluorescence spectrum for a given sample and the EEF method produces a three-dimensional fluorescence profile for each sample. The time required to collect an EEF spectrum is longer as it generates several emission spectra for a given sample when compared to SF mode but the data contain much richer information and could result in better characterizations and quantifications.

Multivariate calibration methods make it possible to relate instrument responses that consist of several predictor variables to a chemical or physical property of a sample. Several classical multivariate calibration methods have been developed (Lindberg et al., 1983; Geladi and Kowalski, 1986; Haaland and Thomas, 1988; Wentzell et al., 1997) in the last couple of decades for the analysis of complex chemical mixtures, and the choice of the most suitable calibration method is very important in order to

generate calibration models with high predictive ability for future samples. In some cases conventional methods may not offer a satisfactory solution to a given problem due to the complexity of the data and it may be necessary to apply some sort of variable selection. There have been many mathematical methods of variable selection and genetic algorithm is one of them offering a fast and effective solution for large-scale problems (Lucasius and Kateman, 1993; Horchner and Kalivas, 1995; Özdemir and Williams, 1999; Özdemir and Öztürk, 2004, 2007).

Inverse least squares (ILS) method is based on the inverse of Beer's Law where concentration of an analyte is modeled as a function of absorbance measurements. Genetic inverse least squares (GILS) is a modified version of the original ILS method in which a small set of wavelengths is selected from a full-spectral data matrix and evolved to an optimum solution using a genetic algorithm (GA) and has been applied to a number of wavelength selection problems. GAs are non-local search and optimization methods that are based upon the principles of natural selection.

In this work, a genetic algorithm-based calibration method called genetic inverse least squares (GILS) was tested with the aim of establishing calibration models that have a high predictive ability for the determination of olive oil adulteration with sunflower oil and corn oil using EEF and SF spectroscopy.

## 50.2 METHODOLOGICAL CONSIDERATIONS

The major drawback of the classical least squares (CLS) method is that all of the interfering species must be known and their concentrations included in the model. This need can be eliminated by using the inverse least squares (ILS) method which uses the inverse of Beer's Law. In the ILS method, concentration of a component is modeled as a function of absorbance measurements. Because modern spectroscopic instruments are very stable and provide excellent signal-to-noise ratios, it is believed that the majority of errors lie in the reference values of the calibration sample, not in the measurement of their spectra. In fact, in many cases the concentration data of calibration set is generated from another analytical technique that already has its inherent errors which might be higher than those of the spectrometer (for example, Kjeldahl protein analysis used to calibrate near infrared spectra).

The ILS model for  $m$  calibration samples with  $n$  wavelengths for each spectrum is described by:

$$\mathbf{C} = \mathbf{A}\mathbf{P} + \mathbf{E}_C \quad (50.1)$$

where  $\mathbf{C}$  is the  $m \times l$  matrix of the component concentrations,  $\mathbf{A}$  is the  $m \times n$  matrix of the calibration spectra,  $\mathbf{P}$  is the  $n \times l$  matrix of the unknown calibration coefficients relating  $l$  component concentrations to the spectral intensities

and  $\mathbf{E}_C$  is the  $m \times l$  matrix of errors in the concentrations not fit by the model. In the calibration step, ILS minimizes the squared sum of the residuals in the concentrations. The biggest advantage of ILS is that Equation 50.1 can be reduced for the analysis of single component at a time since analysis based on an ILS model is invariant with respect to the number of chemical components included in the analysis. The reduced model is given as:

$$\mathbf{c} = \mathbf{A}\mathbf{p} + \mathbf{e}_c \quad (50.2)$$

where  $\mathbf{c}$  is the  $m \times l$  vector of concentrations for the component that is being analyzed,  $\mathbf{p}$  is  $n \times l$  vector of calibration coefficients and  $\mathbf{e}_c$  is the  $m \times l$  vector of concentration residuals not fit by the model. During the calibration step, the least-squares estimate of  $\mathbf{p}$  is:

$$\hat{\mathbf{p}} = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}' \cdot \mathbf{c} \quad (50.3)$$

where  $\hat{\mathbf{p}}$  are the estimated calibration coefficients. Once  $\hat{\mathbf{p}}$  is calculated, the concentration of the analyte of interest can be predicted with the equation below.

$$\hat{c} = \mathbf{a}' \cdot \hat{\mathbf{p}} \quad (50.4)$$

where  $\hat{c}$  is the scalar estimated concentration and  $\mathbf{a}$  is the spectrum of the unknown sample. The ability to predict one component at a time without knowing the concentrations of interfering species has made ILS one of the most frequently used calibration methods.

The major disadvantage of ILS is that the number of wavelengths in the calibration spectra should not be more than the number of calibration samples. This is a big restriction since the number of wavelengths in a spectrum will generally be much more than the number of calibration samples and the selection of wavelengths that provide the best fit for the model is not a trivial process. Several wavelength selection strategies, such as stepwise wavelength selection and all possible combination searches, are available to build an ILS model which fits the data best.

Genetic algorithms (GA) are global search and optimization methods based upon the principles of natural evolution and selection as developed by Darwin. Computationally, the implementation of a typical GA is quite simple and consists of five basic steps including initialization of a gene population, evaluation of the population, selection of the parent genes for breeding and mating, crossover and mutation, and replacing parents with their offspring. These steps have taken their names from the biological foundation of the algorithm.

Genetic inverse least squares (GILS) is an implementation of a GA for selecting wavelengths to build multivariate calibration models with reduced data set. GILS follows the same basic initialize/breed/mutate/evaluate algorithm as other GAs to select a subset of wavelengths but is unique in the way it

encodes genes. A gene is a potential solution to a given problem and the exact form may vary from application to application. Here, the term gene is used to describe the collection of instrumental response at the wavelength range given in the data set. The term ‘population’ is used to describe the collection of individual genes in the current generation.

In the initialization step, the first generation of genes is created randomly with a fixed population size. Although random initialization helps to minimize bias and maximize the number of possible recombinations, GILS is designed to select initial genes in a somewhat biased random fashion in order to start with genes better suited to the problem than those that would be randomly selected. Biasing is done with a correlation coefficient by plotting the predicted results of initial population against the actual component concentrations. The size of the gene pool is a user-defined even number in order to allow breeding of each gene in the population. It is important to note that the larger the population size, the longer the computation time. The number of instrumental responses in a gene is determined randomly between a fixed low limit and high limit. The lower limit was set to 2 in order to allow single point crossover whereas the higher limit was set to eliminate overfitting problems and reduce the computation time. Once the initial gene population is created, the next step is to evaluate and rank the genes using a fitness function, which is the inverse of the standard error of calibration (SEC).

The third step is where the basic principle of natural evolution is put to work for GILS. This step involves the selection of the parent genes from the current population for breeding using a roulette wheel selection method according to their fitness values. The goal is to give a higher chance to those genes with high fitness so that only the best-performing members of the population will survive in the long run and will be able to pass their information to the next generations. Because of the random nature of the roulette wheel selection method, however, genes with low fitness values will also have some chance to be selected. Also, there will be genes that are selected multiple times and some genes will not be selected at all and will be thrown out of the gene pool. After the selection procedure is completed, the selected genes are allowed to mate top-down in pairs whereby the first gene mates with the second gene and the third one with the fourth one and so on as illustrated in the following example:

Parents

$$S_1 = (A_{347}, A_{951}, \#A_{479}, A_{518}) \quad (50.5)$$

$$S_2 = (A_{625}, A_{378}, A_{568}, \#A_{743}, A_{750}, A_{451}, A_{558}, A_{631}, A_{758}) \quad (50.6)$$

The points where the genes are cut for mating are indicated by #.

Offspring

$$S_3 = (A_{347}, A_{951}, A_{743}, A_{750}, A_{451}, A_{558}, A_{631}, A_{758}) \quad (50.7)$$

$$S_4 = (A_{625}, A_{378}, A_{568}, A_{479}, A_{518}) \quad (50.8)$$

where  $A_{347}$  represents the instrument response at the wavelength given in subscript,  $S_1$  and  $S_2$  represent the first and second parent genes and  $S_3$  and  $S_4$  are the corresponding genes for the offspring. Here the first part of  $S_1$  is combined with the second part of the  $S_2$  to give the  $S_3$ , likewise the second part of the  $S_1$  is combined with the first part of the  $S_2$  to give  $S_4$ . This process is called the single point crossover and is common in GILS. Single point crossover will not provide different offspring if both parent genes are identical, which may happen in roulette wheel selection, when both genes are broken at the same point. Also note that mating can increase or decrease the number of instrument responses in the offspring genes. After crossover, the parent genes are replaced by their offspring and the offspring are evaluated. The ranking process is based on their fitness values following the evaluation step. Then the selection for breeding/mating starts all over again. This is repeated until a predefined number of iterations is reached.

Mutation which introduces random deviations into the population was also introduced into the GILS during the mating step at a rate of 1% as is typical in GAs. This is usually done by replacing one of the responses in an existing gene with a randomly selected new one. Mutation allows the GILS to explore the search space and incorporate new material into the genetic population. It helps keep the search moving and can eject GILS from a local minimum on the response surface. However, it is important not to set the mutation rate too high since it may keep the GA from being able to exploit the existing population. Also, the GILS method is an iterative algorithm and therefore there is a high possibility that the method may easily overfit the calibration data so that the predictions for independent sets might be poor. To eliminate possible overfitting problems, cross validation is used in which one spectrum is left out of the calibration set and the model is constructed with  $m - 1$  sample. Then this model is used to predict the concentration of left out sample. This process is continued until all samples are left out at least once in each iteration. As long as the number of spectra in the calibration set is not too large, cross validation is an effective method of eliminating overfitting. If the number of calibration spectra is very large, then the GILS method has the option of half validation approach in which the half of the spectra in the calibration set is used to validate the model in each iteration.

In the end, the gene with the lowest SEC (highest fitness) is selected for the model building and this model is used to predict the concentrations of component being analyzed in the prediction (test) sets. The success of the model in the prediction of the test sets is evaluated using standard

error of prediction (SEP). Because random processes are heavily involved in GILS as in all the GAs, the program has been set to run several times for each component in this study. The best run (i.e. the one generating the lowest SEC for the calibration set and at the same time producing SEPs for prediction sets that are in the same range with the SEC) is subsequently selected for evaluation and further analysis. The termination of the algorithm can be done in many ways. The easiest way is to set a predefined iteration number for the number of breeding/mating cycles.

GILS has some major advantages over classical univariate and multivariate calibration methods. First of all, it is quite simple in terms of the mathematics involved in the model building and prediction steps, but at the same time it has the advantages of the multivariate calibration methods with a reduced data set since it uses the full spectrum to extract genes. By selecting a subset of instrument responses it is able to eliminate non-linearities that might be present in the full spectral region.

The GILS method was applied to the fluorescence spectra of ternary mixtures of olive oil, sunflower oil, and corn oil. The fluorescence spectra of 50 ternary samples of olive oil, corn oil and sunflower oil were measured in EEF and SF mode using a Varian Cary Eclipse spectrophotometer (Varian, Inc. Hansen Way, Palo Alto, CA) equipped with a xenon flash lamp. Olive oil, sunflower oil and corn oil samples were purchased from a local grocery store. The EEF spectra were collected between 320 and 800 nm emission wavelength by exciting the samples with a wavelength increment of 15 nm ( $\Delta\lambda$ ) from 320 to 425 nm. The SF spectra of the samples were recorded between 250 and 750 nm with a  $\Delta\lambda$  of 20 nm. The slit widths of excitation and emission monochromators were set to 5 nm in both EEF and SF modes. All spectra were then transferred to a computer where the data-processing programs were installed. Among the 50 ternary mixtures, 34 of the sample were randomly selected with the exception that the samples with the lowest and highest concentration of each component were

**TABLE 50.1** Percent composition of calibration set used in ternary mixtures of olive oil, corn oil and sunflower oil for both excitation-emission fluorescence (EEF) and synchronous fluorescence (SF) data sets.

S. N.	Olive oil (w/w %)	Corn oil (w/w %)	Sunflower oil (w/w %)	S.N.	Olive oil (w/w %)	Corn oil (w/w %)	Sunflower oil (w/w %)
1	59.87	29.06	11.07	18	68.93	15.01	16.06
2	65.97	26.98	7.05	19	73.89	1.04	25.07
3	61.85	22.08	16.07	20	89.86	4.05	6.09
4	74.86	16.05	9.09	21	73.91	18.00	8.09
5	90.90	3.03	6.07	22	70.96	20.07	8.97
6	88.00	4.00	8.00	23	73.09	11.96	14.95
7	80.93	8.01	11.05	24	97.99	1.01	1.00
8	68.06	22.99	8.95	25	68.86	18.08	13.06
9	71.93	27.02	1.04	26	60.01	14.02	25.97
10	76.91	20.98	2.11	27	72.04	19.02	8.94
11	70.06	29.94	0.00	28	69.93	5.04	25.03
12	70.00	0.00	30.00	29	88.98	2.04	8.98
13	89.90	9.04	1.06	30	84.91	3.00	12.09
14	94.86	2.11	3.03	31	79.09	20.91	0.00
15	66.98	19.94	13.08	32	63.02	5.08	31.90
16	87.04	12.96	0.00	33	62.96	28.03	9.01
17	75.96	0.00	24.04	34	72.97	10.04	17.00

*The same calibration set was used for both excitation-emission fluorescence (EEF) and synchronous fluorescence (SF) data. The total of 34 calibration samples were randomly selected from 50 samples.*

**TABLE 50.2** Percent composition of prediction set used in ternary mixtures of olive oil, corn oil and sunflower oil for both excitation-emission fluorescence (EEF) and synchronous fluorescence (SF) data sets.

S. N.	Olive oil (w/w %)	Corn oil (w/w %)	Sunflower oil (w/w %)	S.N.	Olive oil (w/w %)	Corn oil (w/w %)	Sunflower oil (w/w %)
1	77.99	3.97	18.04	9	80.92	13.00	6.08
2	90.98	2.05	6.96	10	69.93	17.02	13.05
3	59.93	15.07	25.00	11	78.94	18.04	3.01
4	64.05	18.98	16.97	12	76.95	3.98	19.07
5	84.88	5.05	10.07	13	87.94	7.12	4.94
6	82.96	10.94	6.10	14	66.01	8.06	25.94
7	74.00	10.97	15.04	15	78.96	18.97	2.08
8	65.94	12.02	22.04	16	85.93	1.02	13.04

*The same prediction set was used for both excitation-emission fluorescence (EEF) and synchronous fluorescence (SF) data. The total of 16 prediction samples were randomly selected from 50 samples.*

intentionally put in the calibration set in order to construct calibration models as shown in Table 50.1. Table 50.2 shows the concentration profiles of the remaining 16 samples for the validation set as weight percent (Wt-%). The same calibration and validation sets were used in both EEF and SF data. The genetic inverse least squares (GILS) method was written in MATLAB programming language using Matlab 5.3 (MathWorks Inc, Natick, MA).

### 50.3 FEATURES OF MOLECULAR FLUORESCENCE SPECTROSCOPY

Fluorescence spectroscopic measurements can be carried out from simple steady-state emission intensity to quite sophisticated time-resolved measurements. Although fluorescence measurements do not provide detailed structural information, fluorescence spectroscopy is gaining interest in many areas of science for quantitative analysis of complex mixtures with the help of advanced multivariate statistical tools.

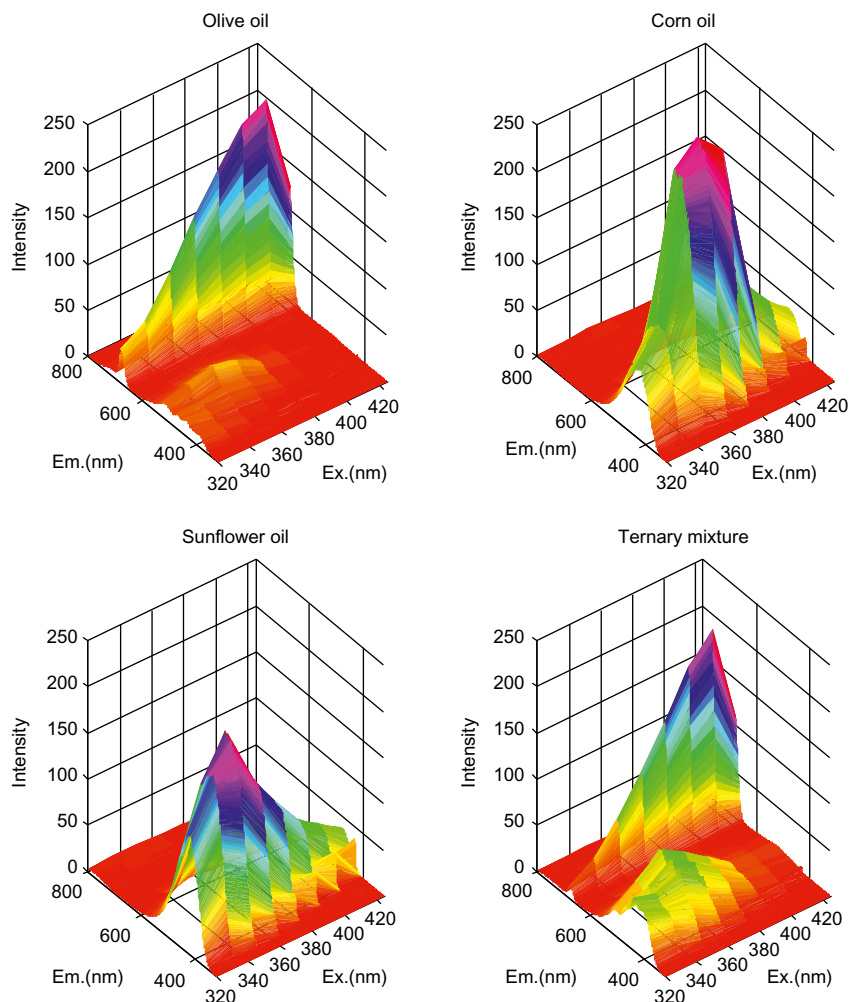
Fluorescence occurs in simple as well as in complex gaseous, liquid, and solid chemical systems. While fluorescence can be observed from almost all molecules with an excitation beam in adequate intensity only a small part of molecules demonstrates fluorescence characteristics which are desirable for analytical purposes. Therefore, fluorescence spectroscopy is less universal than absorption techniques although it is more selective. However, in some applications in terms of its lower detection limits and greater selectivity, fluorescence spectrometry is a preferred technique to molecular absorption spectrometry. Even if many species in a sample are fluorescent, selectivity is

always improved by a suitable choice of excitation and emission wavelengths.

Fluorescence emission is a transition between electronic states of the same multiplicity and involves a singlet-singlet transition. Emission occurs from the ground vibrational level of excited electronic states ( $S_1$  or  $S_2$ ) to various vibrational levels in ground electronic state ( $S_0$ ). Fluorescence usually appears at longer wavelengths than absorption as absorption transitions are higher excited electronic states. For the analysis of multifluorophoric systems the widely used methods are the excitation emission fluorescence (EEF) spectroscopy and synchronous fluorescence (SF) spectroscopy. EEF spectroscopy is a rapid and inexpensive technique and is also known as total fluorescence spectroscopy (TFS) which provides a 'fingerprint' consisting of a 3-D emission/excitation intensity diagram. This 'fingerprint' along with multivariate calibration can be used for the qualitative and quantitative information about the multifluorophores present in the sample. Synchronous fluorescence spectroscopy is a highly sensitive and simple technique. In SF spectroscopy both the excitation and emission monochromators are simultaneously scanned at a constant wavelength interval between emission and excitation wavelengths ( $\Delta\lambda$ ), so that spectral overlaps are reduced and the spectra is simplified.

### 50.4 APPLICATION OF MOLECULAR FLUORESCENCE SPECTROSCOPY TO OLIVE OIL ADULTERATION

Fluorescence spectroscopy has become a popular spectroscopic technique due to its high sensitivity and selectivity.



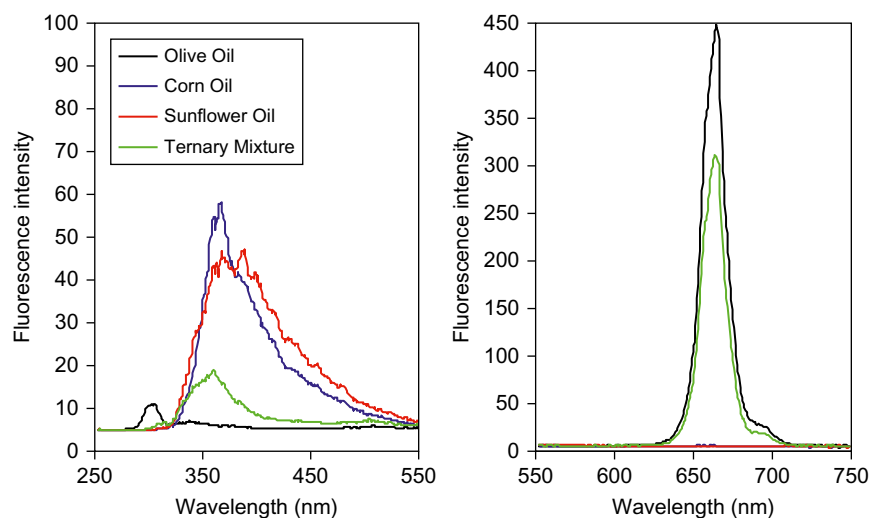
**FIGURE 50.1** Excitation and emission fluorescence spectra show the maximum emission intensities between 320 and 800nm emission wavelengths and between 320 and 425 excitation wavelengths for each component along with their ternary mixture. Olive oil when compared with corn oil and sunflower oil gives quite different emission profile.

Determination of olive oil adulteration with corn and sunflower oils was conducted using fluorescence spectroscopy coupled with genetic multivariate calibration. The EEF spectra of pure olive, sunflower and corn oil and their ternary mixture between the 320 and 800nm emission wavelengths and excitation wavelengths ranging from 320 to 425 nm are shown in Figure 50.1. As seen from the figure, the EEF spectra of corn oil and sunflower oil are very much alike, showing maximum fluorescence emission intensity around 500nm and strongest excitation around 380nm. Pure olive oil has maximum fluorescence emission profile around 700nm with an excitation wavelength of 410nm. In addition, olive oil also gives a weaker emission peak around 500nm which overlaps with corn oil and sunflower emission profile.

Synchronous fluorescence spectra of olive oil, corn oil, and sunflower oil along with their ternary mixture between 250 and 750nm are shown in Figure 50.2. The spectra

were divided into two parts in order to better illustrate the emission profile around 380nm as the intensity at this lower wavelength is about 10 times lower compared to the peak around 670nm which is only seen for olive oil. As can be seen from the figure, corn oil and sunflower oil have very similar synchronous fluorescence emission with a maximum intensity around 350nm, whereas olive oil emission is distinctly different.

A total of 50 ternary mixtures of olive oil, corn oil and sunflower oil were prepared in order to prepare multivariate calibration models. The calibration models were prepared with 34 samples as given in Table 50.1 and then these models were tested with 16 independent prediction samples shown in Table 50.2 for both EEF and SF data. Because of the random nature of the GILS method, the program was set to run 100 times with 30 genes and 50 iterations. Since the GILS program was iterated 50 times in each run, full cross validation was applied during the model building



**FIGURE 50.2** Synchronous fluorescence emission spectra of olive oil, corn oil and sunflower oil along with their ternary mixture between 250 and 750 nm were divided into two parts in order to illustrate high fluorescence emission of olive oil around 670 nm where no significant emission is seen for corn oil and sunflower oil.

**TABLE 50.3** Standard error of calibration (SEC) and standard error of prediction (SEP) results for both excitation-emission fluorescence (EEF) and synchronous fluorescence (SF) data sets.

Data sets	SEC and SEP	Olive oil	Corn oil	Sunflower oil
EEF data set	SEC (w/w %)	0.58	0.51	0.61
	SEP (w/w %)	0.64	0.90	1.07
SF data set	SEC (w/w %)	0.63	0.57	0.81
	SEP (w/w %)	0.64	1.07	1.21

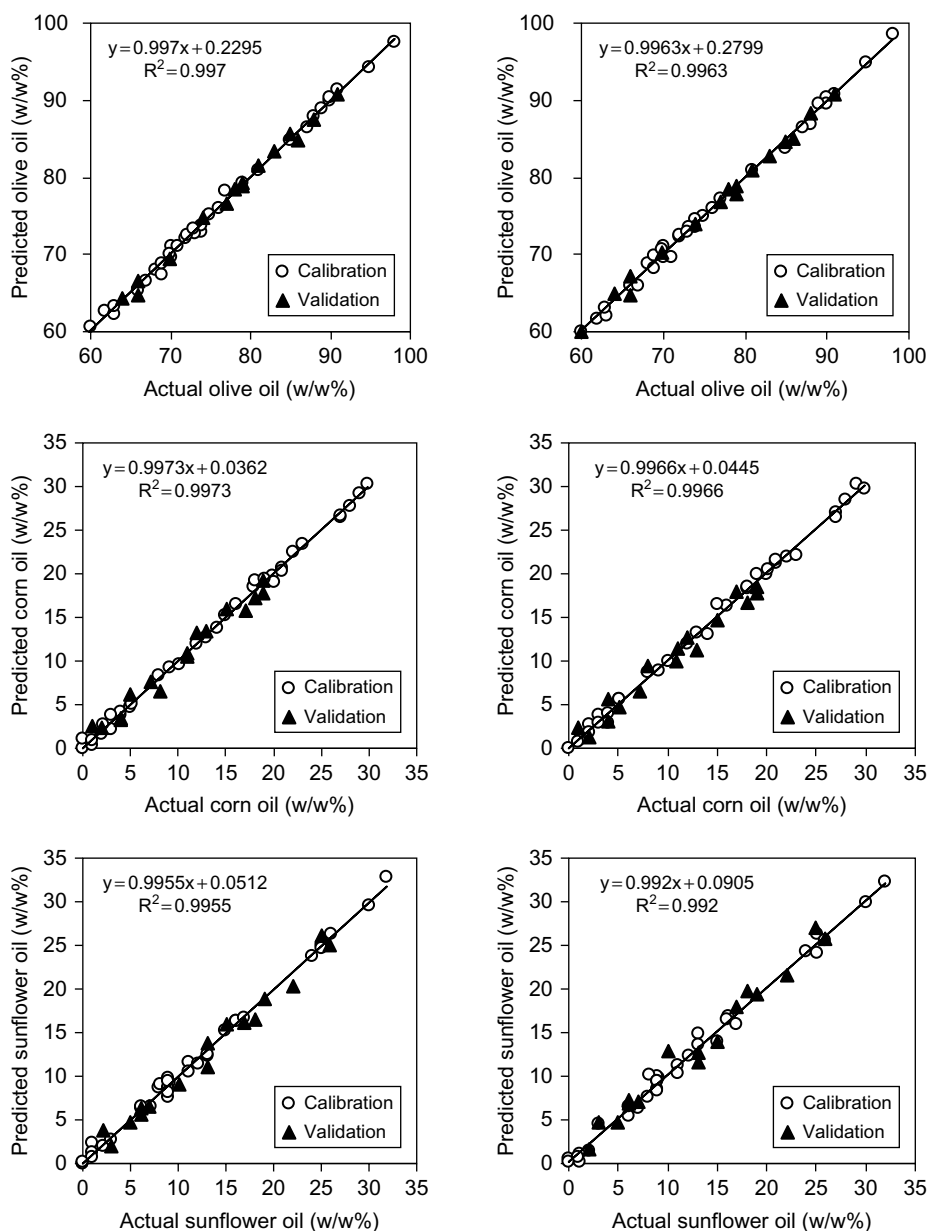
*The SEC and the SEP results for both excitation-emission fluorescence (EEF) and synchronous fluorescence (SF) data sets were given in order to compare the success of GILS generating calibration models that have high predictive ability for the independent prediction sets. The results were ranged between 0.51% (w/w) and 1.21% (w/w).*

step to avoid possible overfitting problems. The standard error of calibration (SEC) and standard error of prediction (SEP) results for both EEF and SF data sets are shown in Table 50.3. As seen in Table 50.3, the SEC and SEP values ranged between 0.51 and 1.27% by mass for both EEF and SF data sets. Considering the fact that any possible olive oil adulteration attempt may include up to 30% or more vegetable oil by volume or by mass, these values seem to be a good prediction for a fast identification. Furthermore, both calibration and prediction results for EEF data were slightly lower than SF data set indicating that better prediction results are obtained with EEF data. This seems

reasonable since the EEF data were obtained at eight different excitation wavelengths from 320 to 425 nm resulting in a richer fluorescence emission profile when compared with SF data. It is also evident from the table that the calibration model generated with both EEF and SF for olive oil gives better prediction results among the three oils used to prepare ternary mixtures. This is no surprise as the fluorescence spectra of olive oil are distinctly different from corn oil and sunflower oil whereas the latter two have very similar fluorescence emission spectra.

The plot of actual versus predicted concentrations for olive oil, corn oil and sunflower oil is illustrated in Figure 50.3 for EEF (on the left) and SF (on the right) data. It is evident that the proposed method is able to predict adulteration of olive oil with corn oil and sunflower oil in a wide dynamic range from 1% to 35% by mass. While the results of olive oil are almost exactly the same for both EEF and SF data sets, the same is not true for corn oil and sunflower oil. The performance of GILS for corn oil and sunflower oil is slightly better for EEF data and this could be due to the partial resolution of their peaks around 500 nm as the excitation wavelength is changed. As a result, it is concluded that both excitation and emission fluorescence spectroscopy in conjunction with multivariate calibration can be used for the fast identification of olive oil adulteration with cheaper substitutes.

Because GILS is a wavelength-selection-based method, it is interesting to observe the distribution of selected wavelengths in multiple runs over the entire full spectral region. The frequency distribution of selected wavelengths in 100 runs for olive oil, corn oil and sunflower oil is illustrated in Figures 50.4 and 50.5 for EEF and SF data, respectively.

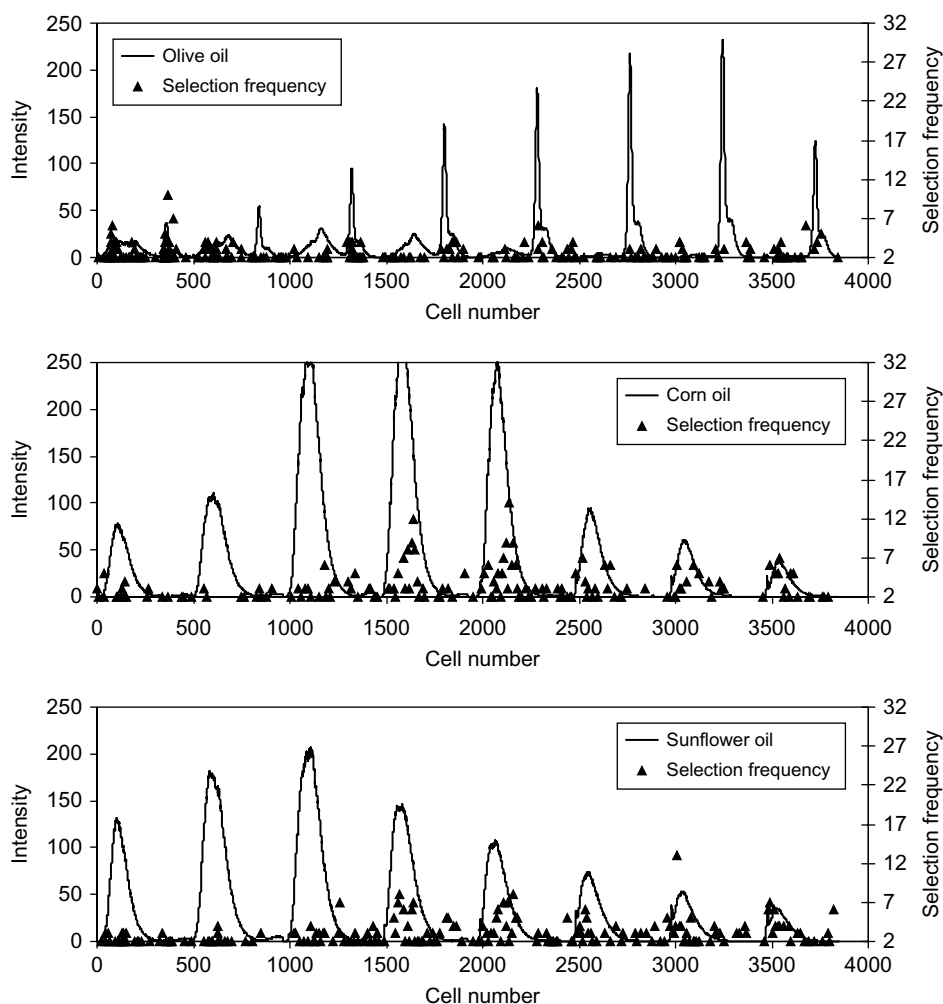


**FIGURE 50.3** Each graph illustrates the success of the GILS method for the construction of calibration models and the validation of these models with independent test sets which are not used in the model building step. In order to make comparison, actual vs. predicted percent olive oil, corn oil, and sunflower oil contents of the ternary mixtures for excitation-emission (on the left) and synchronous fluorescence (on the right) data are given side by side for each component.

When a close examination is done on [Figure 50.4](#), emission spectra of eight different excitation wavelengths were concatenated in order to illustrate both pure component spectra and selection frequency distribution in a simple two-dimensional plot. The most frequently selected wavelengths for olive oil are located in the first two to three excitation wavelengths whereas the most frequently selected wavelengths for corn oil and sunflower oil are concentrated in the fourth and fifth excitation wavelengths and also a few other higher excitation wavelengths. It is also important to note that the selection frequencies were considerably low

in flat baseline portions of the spectra which indicate that the GILS method is able to focus on the information-rich regions of the spectra even though it starts with a completely random selection of wavelengths. In the case of [Figure 50.5](#), an interesting selection frequency profile for olive oil is seen where the most frequently selected wavelengths are located around 300 nm where there is a small peak. The possible reason for this might be the better linearity of the signal at this wavelength region. The intensity scale is reduced to 25 in order to better illustrate this small peak for olive oil. The selection frequency distributions of





**FIGURE 50.4** Three-dimensional data generated from EEF data are concatenated in a way that EEF matrix data were converted to a vector for each sample resulting in a matrix representation of all samples in calibration and validation sets. This is required for GILS in order to generate multivariate calibration models. Distribution of the selected wavelengths for EEF data for olive oil, corn oil and sunflower oil along with their pure component concatenated EEF spectra for eight different excitation wavelengths from 320 to 425 nm shows that GILS selects different excitation and emission wavelengths for each component.

corn oil and sunflower oil samples resulted in expected profiles as shown in the figure. This is a strong indication that the genetic algorithm incorporated into the GILS method is focusing on the regions where most concentration-related information is contained.

## SUMMARY POINTS

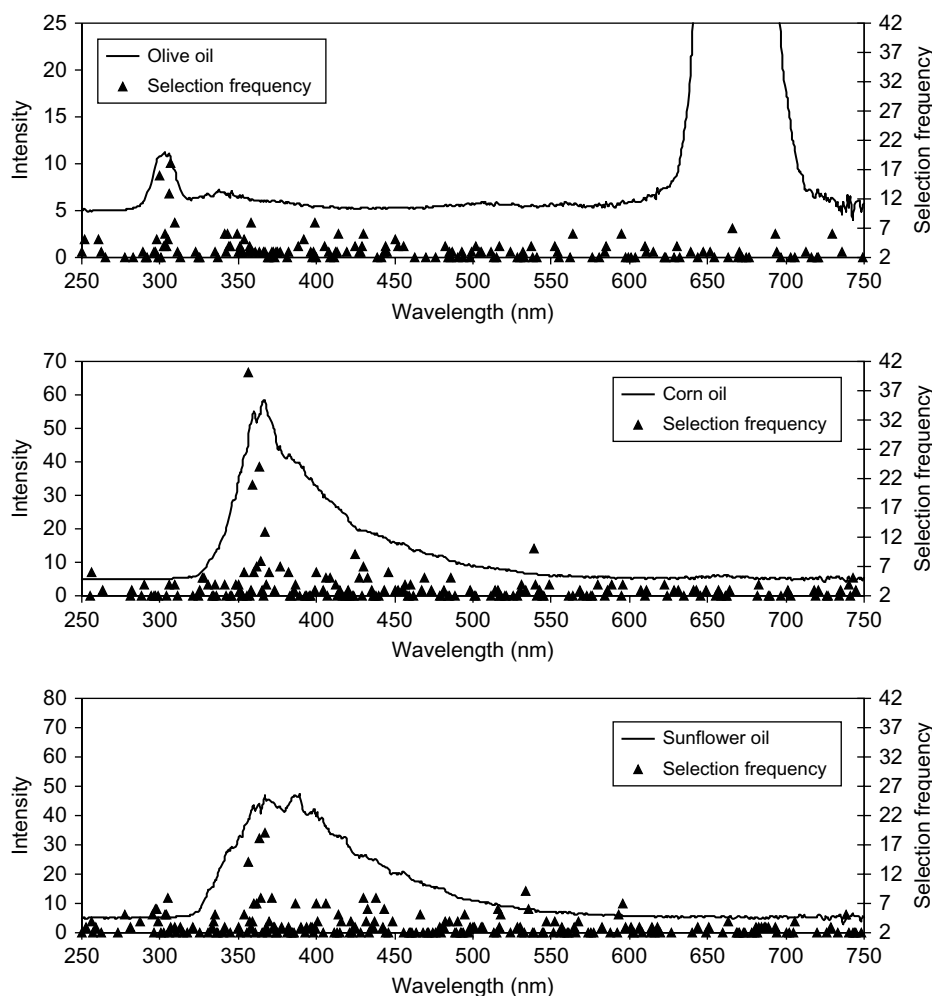
- Adulteration of olive oil with cheaper substitutes such as sunflower and corn oil is a major concern for the public.
- Rapid analysis methods are required for a quick and easy screening of possible adulteration attempts.
- Fluorescence spectroscopy coupled with a genetic algorithm-based multivariate calibration method allows the

determination of olive oil adulteration with sunflower and corn oil.

- The fact that the standard error of prediction values are all below 1.30% (w/w) for the ternary set, fluorescence spectroscopy can be used as a fast screening method for possible olive oil adulteration with cheaper vegetable oils.
- In addition, the genetic algorithm used in the GILS method is able to select and extract the most relevant information to build successful calibration models that has high predictive ability for the independent test samples.

## ACKNOWLEDGMENT

This work was financially supported by the Scientific and Technological Research Council of Turkey (TUBİTAK) through Project No: 107T037.



**FIGURE 50.5** Distribution of the selected wavelengths for SF data set by the GILS method for olive oil, corn oil and sunflower oil along with their pure component SF spectra shows that the method focuses to the regions where corn, olive and sunflower oil gives fluorescence emission. On the other hand, higher selection frequencies for olive oil are seen around 300 nm where a smaller emission peak is given when compared with 670 nm region.

## REFERENCES

- Aparicio, R., Morales, M.T., Alonso, V., 1997. Authentication of European virgin olive oils by their chemical compounds, sensory attributes and consumers' attitudes, *J. Agr. Food Chem.* 45, 1076–1083.
- Christopoulou, E., Lazaraki, M., Komaitis, M., Kaselimis, K., 2004. Effectiveness of determinations of fatty acids and triglycerides for the detection of adulteration of olive oils with vegetable oils. *Food Chem* 84, 463–474.
- Dennis, M.J., 1998. Recent developments in food authentication. *Analyst* 123, 151R–156R.
- Francesca, G., Ricard, B., Joan, F., 2004a. Cluster analysis applied to the exploratory analysis of commercial Spanish olive oils by means of excitation-emission fluorescence spectroscopy. *J. Agric. Food Chem.* 52, 6673–6679.
- Francesca, G., Joan, F., Ricard, F., Xavier, R., 2004b. Application of unfold principal component analysis and parallel factor analysis to the exploratory analysis of olive oils by means of excitation-emission matrix fluorescence spectroscopy. *Anal. Chim. Acta.* 515, 75–85.
- Geladi, P., Kowalski, B.R., 1986. Partial least squares regression: a tutorial. *Anal. Chim. Acta.* 185, 1–17.
- Ghosh, P., Reddy, K.M.M., Sashidhar, R.B., 2005. Quantitative evaluation of sanguinarine as an index of argemone oil adulteration in edible mustard oil by high performance thin layer chromatography. *Food Chem.* 91, 757–764.
- Guimet, F., Fer e, J., Boqué, R., 2005. Rapid detection of olive-pomace oil adulteration in extra virgin olive oils from the protected denomination of origin "Siurana" using excitation-emission fluorescence spectroscopy and three-way methods of analysis. *Anal. Chim. Acta* 544, 143–152.
- Haaland, D.M., Thomas, E.V., 1988. Partial least squares methods for spectral analyses. 1. Relation to other quantitative calibration methods and the extraction of qualitative information. *Anal. Chem.* 60, 1193–1202.
- Hajimahmoodi, M., Vander, H.Y., Sadeghi, N., Jannat, B., Oveisi, M.R., Shahbazian, S., 2005. Gas-chromatographic fatty-acid fingerprints and partial least squares modeling as a basis for the simultaneous determination of edible oil mixtures. *Talanta* 66, 1108–1116.
- Hörchner, U., Kalivas, J.H., 1995. Further investigation on a comparative study of simulated annealing and genetic algorithm for wavelength selection. *Anal. Chim. Acta* 311, 1–13.
- Konstantina, I.P., George, A.M., Constantinos, A.G., 2006. Synchronous fluorescence spectroscopy for quantitative determination of virgin olive oil adulteration with sunflower oil. *Anal. Bioanal. Chem.* 386, 1571–1575.

- Kyriakidis, N.B., Skarkalis, P., 2000. Fluorescence spectra measurement of olive oil and other vegetable oils. *J. AOAC Int.* 83, 1435–1438.
- Lindberg, W., Persson, J.A., Wold, S., 1983. Partial least-squares method for spectrofluorimetric analysis of mixtures of humic acid and lignin sulfonate. *Anal. Chem.* 55, 643–648.
- Lucasius, C.B., Kateman, G., 1993. Understanding and using genetic algorithms. Part 1. Concepts, properties and context. *Chem. Intell. Lab. Syst.* 19, 1–33.
- Özdemir, D., Öztürk, B., 2007. Near infrared spectroscopic determination of olive oil adulteration with sunflower and corn oil. *J. Food Drug Anal.* 15 (1), 40–47.
- Özdemir, D., Öztürk, B., 2004. Genetic multivariate calibration methods for near infrared (NIR) spectroscopic determination of complex mixtures. *Turk. J. Chem.* 28, 497–514.
- Özdemir, D., Williams, R.R., 1999. Multi-instrument calibration in UV-visible spectroscopy using genetic regression. *Appl. Spectrosc.* 53, 210–217.
- Papadopoulos, K., Triantis, T., Tzikis, C.H., Nikokavoura, A., Dimotikali, D., 2002. Investigations of the adulteration of extra virgin olive oils with seed oils using their weak chemiluminescence. *Anal. Chim. Acta* 464, 135–140.
- Schulz, H., Quilitzsch, R., Krüger, H., 2003. Rapid evaluation and quantitative analysis of thyme, oregano and chamomile essential oils by ATR-IR and NIR spectroscopy. *J. Mol. Struct.* 661, 299–306.
- Tay, A., Singh, R.K., Krishnan, S.S., Gore, J.P., 2002. Authentication of olive oil adulterated with vegetable oils using Fourier transform infrared spectroscopy. *Lebensm.-Wiss. u.-Technol.* 35, 99–103.
- Wentzell, P.D., Andrews, D.T., Kowalski, B.R., 1997. Maximum likelihood multivariate calibration. *Anal. Chem.* 69, 2299–2311.
- Wenzla, T., Prettnerb, E., Schweigerb, K., Wagnerc, F.S., 2002. An improved method to discover adulteration of Styrian pumpkin seed oil. *J. Biochem. Bioph. Meth.* 53, 193–202.