

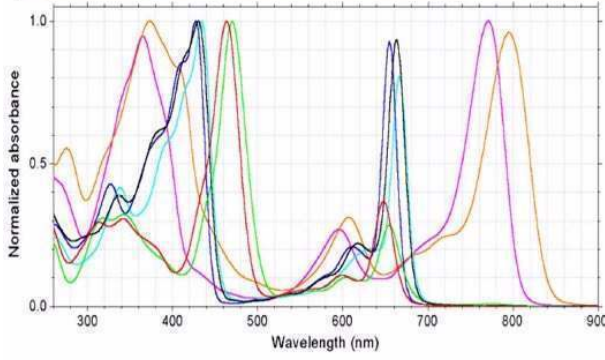
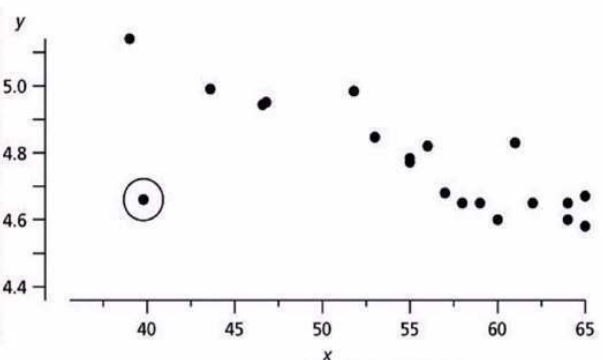


SCHOOL OF CHEMICAL ENGINEERING
 SECTOR OF COMPOSITION AND DEVELOPMENT OF INDUSTRIAL PROCESSES
 LABORATORY OF FUELS AND LUBRICANTS TECHNOLOGY

DEGREE THESIS:

“APPLICATION OF SPECTROSCOPY TECHNIQUES FOR PREDICTING PROPERTIES OF GASOLINE AND DIESEL”

UNL2000		GASOLINE UNLEADED 95 BOM		FINISHED TANK	
17.05.2014	06:49:30	2401A	0.7200	0.7750	4.728
0002 A/A	DEBITMATOR		15.000	0.7292	
0010	DENSITY AT 15 C		0.7200	0.7292	
0020	Research Octane Number		85.0	86.0	
0030	MCH		1.000	1.000	
0040	Induction Period		0.000	0.001	
0060	Lead in Gasoline		0.0	0.0	
0062	SULFUR, ppm		10.0	7.4	
0070	Existent Gum		0	2	
0080	DVPE		45.0	60.0	
0080	DVPE		0.00	0.87	
0100	Evaporated at 70 C		20.0	32.0	
0110	Evaporated at 100 C		48.0	61.0	
0112	DISTIL. EVAP 150cc		75.0	93.0	
0130	Distillation, FBI		100.0	186.0	
0140	Distillation Residue		0.0	2.2	
0160	AROMATICS		0.0	16.4	
0180	Oxygenates Content		0.00	5.60	
0200	Copper Corrosion		0.0	1A	
0210	Oxygen		0.0	1.1	
0230	OLEFINS		8.0	18.0	
0240	DENS AT 15 C UPPER		0.0000	0.7292	
0240	DENS AT 15 C MED		0.0000	0.7292	
0240	DENS AT 15 C BOT		0.0000	0.7292	
0260	DVPE UPPER		45.0	60.0	
0270	DVPE MED		45.0	60.0	
0280	DVPE BOT		45.0	60.0	
UNL2000					
17.05.2014	15:06:27	2401A	0.7200	0.7750	4.843
0062 A/A	DEBITMATOR		15.000	0.7296	
0010	DENSITY AT 15 C		0.7200	0.7296	



JOHN - HECTOR CHALOULOS
 Supervising Professor : Karonis Dimitrios

Acknowledgements

The current work entitled "Application of spectroscopy techniques for predicting properties of gasoline and diesel" was carried out during the academic year 2013-2014, as part of the educational process in the Laboratory of Technology of Fuels and Lubricants of the School of Chemical Engineering of NTUA.

Particularly important factor for the completion of my work was the favorable climate of cooperation with all laboratory staff that helped me with great eagerness for my adaptation with the environment. At this point in particular I would like to thank the head of the subject, Professor of Chemical Engineering, Mr. Karonis Dimitrios for the confidence he showed me by assigning this work to me and his decisive advices for any arising issue based on his experience and knowledge.

I also must pay many thanks to Professors Mr. F. Zannikos and Mr. E. Lois who together with Mr. Karonis trusted me in handling valuable equipment such as the Eraspec instrument.

Very important was the contribution of Mr. Kotsokolou and Mr. Chatzigakis, responsible in the refineries of Motor Oil Hellas and of EL.PE respectively, who provided about 50 samples with certified properties for the purposes of this work.

Finally I must thank those responsible of the technical section of Eralytics, Dr. Philip Jordan, Dr. Christoph Wagner and Dr. Christensson Niklas, who solved any questions that arose with regards to handling the instrument and its maintenance, with their e-mails.

*John Chaloulos
Athens, October 2014*

Contents

1. Introduction.....	7
1.1 General.....	7
1.2 The workflow	7
2. Theoretical background.....	8
2.1 Introduction to spectroscopy.....	8
2.2 Infrared spectroscopy.....	12
2.3 FTIR.....	14
2.4 Analysis of general multicomponent mixtures.....	21
3. Methods of predicting properties / calibration.....	24
3.1 PCR method.....	24
3.2 Cluster Analysis method.....	26
3.3 MLR method.....	28
4. Application of FTIR in organic molecules.....	28
4.1 Introduction.....	28
4.2 Aliphatic hydrocarbons.....	28
4.3 Aromatic components.....	30
4.4 Oxygen containing components.....	32
4.5 Tables of various wavenumbers.....	33
5. Fuel analysis using FTIR.....	36
5.1 Gasoline analysis.....	37
5.2 Diesel analysis.....	39
6. Database.....	41
6.1 Libraries.....	41
6.2 Creation of a new library.....	51
6.3 Refinery protocols.....	53
7. The statistical methods in the instrument.....	61
7.1 Results for the calculated CA properties.....	61
7.2 Results for the calculated MLR properties.....	65

Contents (cont.)

8. Exporting spectra from the device.....	77
8.1 Identification techniques	77
8.2 Sample types.....	78
8.3 Display of spectra.....	79
8.4 Comparison of gasoline spectra	85
8.5 Presentation of diesel spectra	88
8.6 Comparison of diesel spectra	91
9. Conclusions / Results.....	96
10. Bibliography.....	100

Abstract

Fuel is the main pillar of the industry and transport. Therefore the growing need for continuous monitoring fuel has set the bar in measurement precision too high. To face this very problem, specific spectroscopy techniques have been developed, which in combination with particular data processing can provide quick and immediate solutions, predicting gasoline and diesel properties with fairly good approximation.

Gasoline and diesel are complicated samples and consist of many compounds. They consist of organic compounds with, as all molecules, various characteristics in their structure and consequently in the bonds connecting the atoms in their molecules. For example, some of them are changes in the angle of the planes formed by methyl groups, or something simpler, the extent and oscillation of C-H bond.

Infrared light is divided into three regions, near (proximal), mid and far (distal) IR. In the current work the mid infrared extending from 400 to 4000 cm^{-1} is studied. When the mid IR light is incident on fuel samples under specific conditions, interacts with them and light absorption occurs. This absorption is due to oscillations between the atomic bonds of the sample molecules. This process is the Infrared Spectroscopy. If this process is combined with Fourier transformation of the absorption signal, then it can be imprinted by means of mathematical relationships to create a spectrum. The part of the spectrum between 500 to 1600 cm^{-1} is also called fingerprint section, because it is characteristic of each substance.

To obtain a spectrum with FTIR a particular arrangement is needed which together with the interferometer form a spectrophotometer. Such an integrated device is the Eraspec apparatus that was used for the experimental measurements. By applying the Beer Lambert Law in the function of the interferometer and the light beam, and by converting the analog signal to a digital one, the spectrum is obtained on the display of the device. In order for the spectrum to be useful, it must give some information to the user. Also the device in order to be able to determine and predict the properties of the fuel, its software must be supplied with other samples spectra of known properties, so as to make comparisons with the unknown samples.

The Eraspec instrument is used for measurements on gasoline and diesel. From the manufacturer data are set for both types of fuel, i.e. it is calibrated with hundreds of samples of known properties loaded in its libraries. However the manufacturer is recommending further calibration with samples of known properties from the local refineries, namely Motor Oil Hellas and Hellenic Petroleum (EL.PE.). This is suggested because the unknown spectra that will be encountered by the analyst will certainly be closer to the standards of the above refineries, and hence the results will be more accurate. Furthermore the supply of crude oil to the refineries may change so that the standards change again slightly in relation with the samples under testing. This might seem difficult at first glance, however, it could be proven much useful and easygoing.

The identification made by Eraspec instrument in its software, is not a simple comparison of spectra, but based on mathematical statistical methods the best available prediction of properties is brought about. The methods used by the device are the MLR (Multi Linear Regression) and the Cluster Analysis (CA). These methods can give aptly conclusions about the properties of unknown samples of the same kind, from a spectral portion whose

properties are known. Moreover, the device displays the results in relation to convening the Mahalanobis result and distance, which in a way is a measure of the method accuracy.

This was the main objective of the work: The enrichment of the Eraspec instrument database with spectra from the Greek refineries. In total 49 samples were taken with measured properties. These samples were carried to the laboratory and within a short time were entered in new libraries of the instrument, made by the user.

The procedure was the same for diesel and gasoline. The user after the measurement takes the spectrum, types the properties from the analysis reports and stores them in the library. The properties required by the library are: RON, MON aromatics, olefins, IBP, T10, T50, T90, FBP for gasoline and Cetane number, Cetane index, Cetane improver, aromatics, PNA, IBP, T10, T50, T85, T90, T95, FBP for diesel. From these parameters the Eraspec instrument calculates all the other properties. Obviously the user depending on his needs may change the peak or the MLR model, with regards to the properties he is interested in.

At the end of the experimental procedure, several spectra were obtained from the device via a USB. Actually, absorption values were obtained and then using Excel with processing and mapping to specific wavenumbers, there was a spectrum received. Specifically 8 gasoline spectra were obtained, seven from refinery mixed fractions and one from gasoline station 95 RON, while for diesel the spectra taken were from 6 samples, of which five refinery fractions from different processing units (hydrocracker, hydrotreater, synthetic etc.). Then, these spectra were compared to each other and their differences were identified. Very characteristic was that at very strong or faint absorptions, the values showed error. This has to do with the signal to noise ratio. The more light a substance absorbs the less light reaches the detector, resulting in incorrect recording. The same applies in the reverse case. The most complete record of the spectra occurs at the part known as the fingerprint part. The user can guess and compare the spectra manually, following certain rules.

The techniques based on IR spectroscopy are the most appropriate for predicting fuel properties. The benefits are based on its three features: analytic, spectroscopic and instrumental. These techniques are based on the movement of fundamental particles, which means that detections are done almost at the molecular level and therefore with very high accuracy in results. In addition, with the development of computers and the very fast data processing, valuable time is saved by analysts. The statistical methods very rapidly converge to a result or the estimation of many properties can take place by measuring a few of them with the instrument. Generally the method as a whole process is very quick and cheap, because among others, does not require specialized personnel. A single member of the laboratory, knowing the basics can begin to conduct typical measurements. Therefore the method is also called "the one button method".

1. Introduction

1.1 General

Today when developments in industry run rapidly, there is constantly a growing demand for rapid quality control and analysis of fuel. This has triggered the growth of portable analytical equipment which enables in situ analyses, comparable in performance with those in laboratories. Their design makes it easy for non-expert users to carry out complex analyses, while the extent of the analysis is significantly reduced taking into account that the measurement takes place outside the laboratory.

Such an analytical instrument is the Eraspec of the Austrian firm Eralytics GmbH. This work dealt with the calibration of this device, using samples from the Greek refineries. It is well known that all modern analytical instruments in order to function effectively they must be calibrated, i.e. to have some starting points (standards), on the basis of which they may predict properties or measure unknown samples. For example, a precision balance needs a calibration at regular intervals. That can be accomplished by weighing objects of known mass. This way the user with his action basically "teaches" the instrument, how much is one, ten or one hundred grams, depending on how the calibration takes place. In other words the instrument "understands" the quantity "one gram" so that when an object of unknown mass is placed on it, through simple mathematics it calculates how many "one gram" are equivalent with this object and based on this, exports the result. All this is based on an initial convention which the international community recognizes as acceptable mainly for communication purposes. However the case of the precision balance is simple, since mass is constant in all countries of the world and the weight has no significant differences, except in rare cases, when a laboratory is located at high altitude. Moreover the calibration weights are not altered with time pass.

Unfortunately regarding the calibration of the spectrophotometer things are much more complicated. The infrared radiation by itself presents difficulties in handling the interpretation of the spectra generated when interacting with fuel. Fuels by nature (gasoline and diesel) are complex mixtures of hydrocarbons so that their spectra are complex too. Therefore it makes sense that their interpretation is very difficult. This requires the development of complex mathematical models to convene to a result through comparison with the standard spectra. Moreover, to obtain standard samples for the calibration is a difficult process as in Greece it is strictly forbidden to export any fuel out of the refineries due to tax purposes. This fact, in combination with the limited lifetime of the samples after their receipt, makes the calibration of Eraspec a particularly laborious process.

1.2 The workflow

To make the calibration first we should obtain samples from the Greek refineries. This process proved to be very complicated because as mentioned above it is strictly forbidden to export the fuel out of the refineries, even though it is small quantities. Even if the fuel sampling was made by the head of the laboratory, the samples should exit the refinery unseen. Moreover the managers in the refineries as well as the auditors are too busy so that a delay in the preparation of the samples seems to be unavoidable. It should be noted here that

not only sampling but also matching each sample with the corresponding certificate specifying its properties must be carefully controlled. In the main part of the paper there is an extensive presentation of the certificates. It is obvious that not all the samples were taken in one visit but little by little at intervals, always according to the program of enrichment of the NTUA libraries. This fact, i.e. that the samples are not all taken at the same period, is positive in the sense that, depending on the tank of the fuel or on the origin of the work or the composition of the produced gasoline and diesel, they are quite different.

With the aim to take full advantage of the samples it was necessary the student to be trained in Eraspec software and the environment of his work. In theoretical level this was done with the assistance of a technician who was trained in the factory where the instrument is produced and with intensive study of the instruction manual of the instrument. It should be noted here that in case of any serious problem concerning the settings or any blocking of the software (bug) the student could contact the person of the factory responsible for software issues in Austria. On a practical level, the first test measurements were performed with gasoline and diesel from a gas station. It should be noted that the student was trained extensively in all functions and accustomed to all the sensitivities of the instrument software and not only in the operation regarding the calibration. Thus it was necessary to study the whole theory of IR spectroscopy, which is the basic principle of the instrument function. More specifically, in the work it is clarified why FTIR is the focal point of the instrument, and also all the benefits of this method for the specific application are presented. A step further, besides the FTIR principle, the export and matching of the spectrum in terms of quality and quantity of data, there also exists the issue of identification. Obviously no sample is the same to some other and an exact match between a standard and a sample under measurement is also impossible. Therefore, various methods linear and non linear using mathematical models that offset existing spectra with their properties and predict the properties of the test samples with high accuracy, have been developed.

In purely experimental part, it is necessary to present and describe the process of a measurement always aiming to enter standards in the libraries of the instrument. Besides the process of measurement is the same whether the operator intends to enter a standard or not. Together with the description of the measurement process, diverse details associated directly with it and with the theory as well as other applications dealing with tweaks for better handling of the instrument, are presented. Finally the measurements take place, and the spectra of various fractions of gasoline and diesel are obtained.

The experimental procedure in essence is the process of measuring the samples and entering their registered properties into the libraries.

2. Theoretical background

2.1 Introduction to Spectroscopy

Infrared Spectroscopy is based on the interaction of matter with the infrared light. It is based on the fact that the molecules absorb specific frequencies, characteristic of their structure. These absorptions are characteristic frequencies. For example, the frequency of the absorbed radiation is characteristic of the energy transfer from the bond or from the vibrating atom.

This energy is defined by the shape of the active surfaces of the molecule, by the mass of the atom, as well as by the relevant electro-oscillatory couplings. In order for an oscillation pattern in the molecule to be "infraredly active", i.e. to be able to imprint a spectrum, it must be related with changes in the dipole. Thus the permanent dipoles are not imprinted, but only those that change instantaneously.

Proceeding one step further, a molecule can vibrate/oscillate with various ways, and each way is characteristic.

Molecules consisting of N atoms:

- Linear molecules of this type have $3xN-5$ vibration/oscillation degrees of freedom, while
- Non-linear molecules have $3xN-6$ vibration degrees of freedom.

For example water (H_2O) which is a non-linear molecule has $3x3-6 = 3$ degrees of freedom in vibration.

The types of vibration are:

- *Stretching*: change in bond length, such as C-H or C-C.
- *Bending*: change in angle between two bonds, such as the H-C-H angle in methyl group.
- *Rocking*: a change in the angle between a group of atoms such as a methyl group, with the rest of the molecule.
- *Wagging*: a change in the angle between the plane of a group of atoms, as in methyl group, with the rest of the molecule.
- *Twisting*: change in angle between the planes of two groups of atoms such as a change in the angle between two methyl groups.
- *Out-of-plane*: a change in the angle between any of the C-H bonds and the plane defined by the remaining atoms in the molecule of ethylene. Another example is the BF_3 when B atom moves into and out of the plane of fluorine.

Simple diatomic molecules have only a single bond between their atoms and only one oscillation/vibration band. If the test molecule is symmetrical e.g. N_2 , the oscillation band is not imprinted in the IR spectrum. Non symmetrical diatomic molecules e.g. carbon monoxide (CO) absorb in the IR spectrum. More complex molecules have more bonds, and the oscillation/vibration spectrum they give is obviously more complicated. Large molecules for example, have many peaks in their IR spectrum.

2.1.1 Advantages of spectrometric techniques

The superiority of NIR, MIR and RAMAN spectroscopic techniques is big. The techniques are many and applicable in diverse areas. *Table 2.1.1* shows the main characteristics of these techniques. The characteristics are categorized into three categories: analytical, spectroscopic and instrumental. For each of them and each technique, there has been added a small design where it is indicated whether a feature is an advantage, partly advantage or disadvantage. The references where these features are originated from are included in *Table 2.1.1*.

Table 2.1.1
Analytical, Spectroscopic and Instrumental characteristics of the NIR, MIR and Raman spectra

	FEATURES	NIR	MIR	Raman	REFERENCES
ANALYTICAL	Rapid analysis	☺	☺	☺	(Gerrard 1991)
	Simple analysis: concentration vs peak area	☹	☺	☺	(Pelletier & Davis 1996)
	Qualitative and quantitative analysis	☺	☺	☺	(Baranska 1987)
	None or reduced sample preparation	☺	☺	☺	(Gerrard 1991)
	Standard calibration and transfer procedures	☺	☺	☹	
	Validated methods (IUPAC, AOCS, ISO)	☹	☺	☹	
	Sample form and sample size	☺	☺	☺	
	No use of pollutant solvents	☺	☺	☺	
	Economic method (e.g. ● saving of labour ● reduction of samples to analyse)	☺	☺	☺	(Scotter 1997; Murray 1999)
	Direct, non-invasive and non destructive in-situ analysis	☺	☺	☺	(Gerrard 1991) (Hendra 1997)
SPECTROSCOPIC	Emission process	na	na	☺	(Vickers & Mann 1991)
	Intensity proportional to concentration	na	na	☺	(Gerrard 1991)
	Well-resolved spectra	☹	☺	☺	(Gerrard 1991; Wilson 1991)
	High resolution spectra	☺	☺	☺	(Levin & Lewis 1990)
	Precise spectral frequency measurement	☺	☺	☺	(Levin & Lewis 1990)
	Non-polar group information	☹	☹	☺	(Terpinski 1987)
	More significant information	☺	☺	☺	(Schradler 1996; Olinga & Siesler, 2000)
	Structural selectivity	☹	☺	☺	(Olinga & Siesler, 2000)
	Signal intensity	☺	☺	☹	
	Trace element determination	☹	☺	☺	(Li-Chan 1996)
	Fluorescence-free spectra	na	na	☺	(Chase 1987)
	Spectra insensitive to temperature	☹	☹	☺	(Pelletier & Davis 1996)
Sample presentation technique	☺	☺	☺		
INSTRUMENTAL	Push-button instrumentation	☺	☺	☺	(van de Voort 1994)
	Works well with aqueous samples	☺	☹	☺	(Vickers & Mann 1991)
	Suitable for use at-, on- and in-line process	☺	☺	☺	(Scotter 1997; Hendra 1997)
	Compatible with suitable fibre optics	☺	☺	☺	(Lewis <i>et al.</i> 1988; Gerrard 1991; Keller <i>et al.</i> 1993; Olinga & Siesler, 2000)
	Process monitoring	☺	☺	☺	(Olinga & Siesler, 2000; Ozaki <i>et al.</i> 1992)
	Microscope	☺	☺	☺	(Williams & Batchelder 1994)
	Camera (Multispectral measurement)	☺	☺	☹	

2.1.2 Analytical characteristics

The analytical characteristics of the near and mid IR spectroscopy have some features that make them catalytic in the area of laboratory analyses. One important factor is time. The example of determining the concentration of iodine in a solution is used here. The time required for this determination using a conventional technique, e.g. a colorimetric method, is 30 to 60 minutes, whereas with a spectroscopic technique from those mentioned above it is at most 5 minutes. Generally the duration of analysis in industrial samples is very important because products like food or fuel may be spoiled, while the production still goes on and new batches should be checked.

Another advantage of spectroscopic methods is that no sample preparation is required. The form of the sample and its quantity are not prohibitive.

Spectroscopic techniques are relatively inexpensive since they do not require time or the presence of catalytic substances. More specifically spectroscopic methods could be particularly useful for research programs since they allow the selection and consequently the reduction in the number of samples for analysis. The spectra allow the researcher to explore a number of samples before deciding what he must measure through the use of a conventional technique. As mentioned by Ian Murray "Within a closed population of 1000 samples it is best to use the spectral solution for 100 representative samples for reference analysis, since after proceeding to calibration it will be easy to predict the rest 900 of them". So the investigator saves 90% of laboratory work.

Additional advantages of spectroscopic techniques are 1) structural analysis 2) qualitative and quantitative determination of elements 3) direct and in situ analysis. From the perspective of analysts, disadvantages of these methods are that they need processing via specialized software, they are difficult to calibrate, and last, few are the official certified methods using such techniques. This is not precisely a disadvantage in terms of efficacy because the instrument is able to detect the suspicious samples expeditiously and then send it to the central labs for a more thorough analysis.

2.1.3 Spectroscopic characteristics

One of the main features of these techniques is that the generated spectra are of high resolution and through proper mathematical processing of the spectrum the exact frequency is obtained. In the mid infrared spectroscopy (and Raman) the peaks are analyzed and their absorptions are assigned to specific chemical species. There is a small error in the processing of spectra but the quantitative analyses are still acceptable. Instead, in the near IR spectrum a series of inter-overlapping frequencies makes difficult to distinguish the chemical species. Another common feature is that they exhibit high structural selectivity and contain further information necessary for the clarification of the structure. Furthermore, these two techniques can provide information about the phase, the crystallinity or the configuration of the unknown sample.

Raman and MIR spectroscopy can be applied to identify marker substances. Especially for these applications most suitable is Raman and the most famous example is the measurement of carotenoids (in ppm) in food production. The carotenoid peaks are observed in the Raman spectrum. A corresponding example of the MIR is the prediction of organochlorides at low levels, in water.

2.1.4 Instrumental characteristics

One of the most interesting features in instrumental spectroscopy is to develop and conduct measurements with one button. "The one button measurement". This is very important because most measurements can be made by beginners who have little familiarity with the subject. Hiring or even training qualified personnel is not necessary.

2.2 Infrared spectroscopy

The infrared (IR) spectroscopy is the spectroscopy involved with the infrared portion of the electromagnetic spectrum, and substantially it is light having a longer wavelength and lower frequency than visible spectrum. The infrared portion of the spectrum of the electromagnetic radiation is divided into three parts: the proximal, the middle and the distal infrared, depending on the distance from the visible portion of the spectrum. (N-IR, Mid-IR, F-IR respectively) The instrument in question is related with spectra falling in the middle part of the infrared. Regarding the middle section it is approximately 4000-400 cm^{-1} of wave number.

$$\text{Wavenumber: } k = \frac{2 \cdot \pi}{\lambda}$$

2.2.1 The middle portion of the infrared radiation

The middle portion of the infrared radiation (4000-400 cm^{-1}) can be divided approximately into four ranges, and the nature of a group of frequencies may generally be defined by the area in which they are localized. Generally the ranges are as follows: X-H bond area (4000-2500 cm^{-1}), the region of the triple bond (2500-1500 cm^{-1}) and the fingerprint region (1500-600 cm^{-1}).

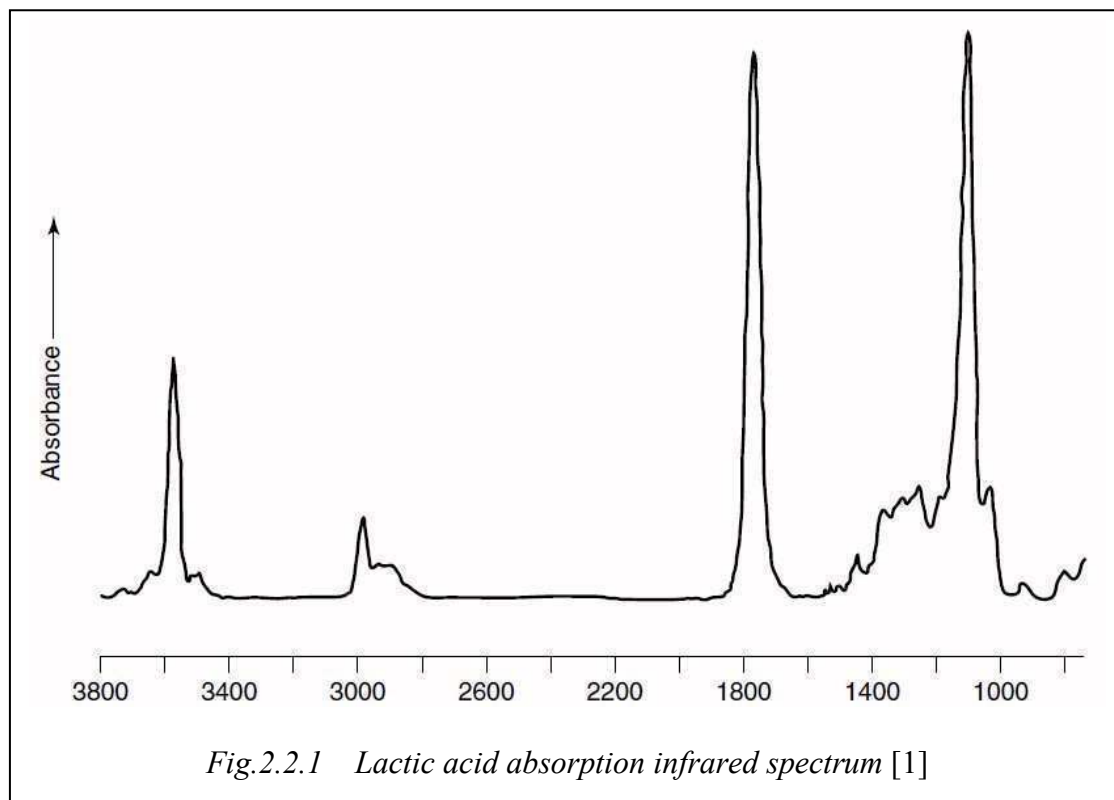
The fundamental vibrations in the region (4000-2500 cm^{-1}) are generally due to the (active) extents of the bonds O-H, C-H and N-H. The extent of O-H bond produces a wide frequency occurring in 3700-3600 cm^{-1} . Compared with this, also the N-H bond is usually observed at 3400 and 3300 cm^{-1} . This absorption is generally much sharper (clearer) than that of the O-H bond, and therefore it can be distinguished. The C-H bond extents from aliphatic compounds are found in the range 3000-2850 cm^{-1} . If the C-H bond is adjacent to a double bond or an aromatic ring, the wavenumber of the C-H extent increases and absorbs between 3100 and 3000 cm^{-1} .

The absorptions of the triple bond extents, fall in the range 2500-2000 cm^{-1} , due to the continuous high strength of the bond. $\text{C}\equiv\text{C}$ bonds absorb at values between 2300 and 2050 cm^{-1} , while in parallel the nitrile group ($\text{C}\equiv\text{N}$) absorbs between 2300 and 2200 cm^{-1} . These groups can be distinguished from each other since the extent of the $\text{C}\equiv\text{C}$ bond is usually very weak, while that of the $\text{C}\equiv\text{N}$ is moderate. These are the most common absorptions in this area, but the researcher may face several extensive absorptions of the X-H type, wherein X is

an atom with larger mass number such as phosphorus or silicon. These absorptions mainly occur approximately near 2400 and 2200 cm^{-1} .

The main frequency bands at the range of $2000\text{-}1500\text{ cm}^{-1}$ are due to the extents of C=C and C=O bonds. The carbonyl group bond extent is one of the easiest absorptions for someone to identify in an infrared spectrum. Usually it is the most intensive range in the spectrum and depending on the type of the C=O bond, it falls in the region between 1830 and 1650 cm^{-1} . It should also be noted that metallic carbonyls can absorb even above 2000 cm^{-1} . The C=C bond is much weaker and is distinguished at around 1650 cm^{-1} . Usually though, this range is not distinct due to symmetry or instantaneous dipoles. The C=N bond occurs in this region and usually is stronger.

So far it has been assumed that each frequency range in an infrared spectrum can be attributed to a specific deformation of the molecule, the movement of a group of atoms, tilting, bending and extent of a particular bond. This performance is possible for several frequency ranges, especially for extent vibrations from multiple bonds showing "good behavior", which means easily distinguishable. However, many vibrations do not operate this way and can vary in hundreds of wavenumbers, even for similar molecules. This happens in most torsional and skeletal vibrations, which absorb in the region of $1500\text{-}650\text{ cm}^{-1}$. The spectrum of a molecule may have a hundred or more regions of absorption, but it is not necessary identifying all of them. The spectrum is considered as the "fingerprint" of the molecule, and therefore this region is referred to as the footprint region.



2.3 FTIR

The infrared spectrum of a sample is obtained when a beam of infrared light is incident on the sample. When the frequency of the beam is the same as the frequency of oscillation / vibration of a bond, there occurs absorption. Processing this beam of light reveals how much energy is absorbed and at what frequency (or wavelength). In the past this was done with a monochromator. In more modern applications, the entire wavelength is measured integer using Fourier transformation devices and thus an absorption spectrum is generated through the process described in the following paragraphs. Following the description of the infrared (IR) the term "FT" (Fourier transform) must be clarified. FTIR spectroscopy is based on the idea of the interference of a radiation between two rays for producing an **interferogram**. The latter is a signal generated as a function of alteration of the traveled length between the two beams. The two fields that take values, i.e. distance and frequency, are interconvertible via the mathematical method of Fourier transformation. The basic equation of Fourier transformation related to the intensity received by the detector, $I(\delta)$, in the spectral dynamic density at a specific wavenumber, $\bar{\nu}$, given by $B(\bar{\nu})$, is:

$$I(\delta) = \int_0^{+\infty} B(\bar{\nu}) \cos(2\pi\bar{\nu}\delta) d\bar{\nu}$$

which is one half of the cosine of the pair of Fourier transformation with the other half being:

$$B(\bar{\nu}) = \int_{-\infty}^{+\infty} I(\delta) \cos(2\pi\bar{\nu}\delta) d\delta$$

These two equations are interconvertible and are known as Fourier transformation pair. The first shows the change in dynamic density as a function of the difference in path (mirror) which is a patent of interference. The second shows the change in intensity as a function of the wavenumber. Each one may be converted to the other by means of the mathematical method of Fourier transformation.

The basic experiment to obtain an FTIR spectrum is to produce an interferogram with the use of a sample and a second interferogram without sample in the beam, and then convert the interferograms to spectra (a) from the source with absorptions from the sample and (b) from the source without absorptions from the sample. The frequencies of the first and the second correspond to a dual beam scattered spectrum.

The great advantage over the use of routine use of the mid infrared came with a new mathematical method (or algorithm) designed for fast Fourier transformation (FFT). That was combined with developments in the processing software that made these calculations to be carried out very quickly.

The general flow diagram of an FTIR spectrophotometer qualitatively is shown in *Fig.2.2.2* and schematically in *Fig.2.2.3*. The detailed process is explained below. It should also be noted that in the flowchart the suffix "computer" does not necessarily mean a computer but

generally computer processing as in the case of Eraspec where the process is made through the instrument's software.

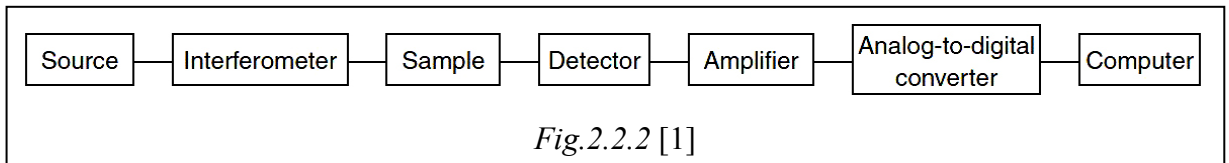


Fig.2.2.2 [1]

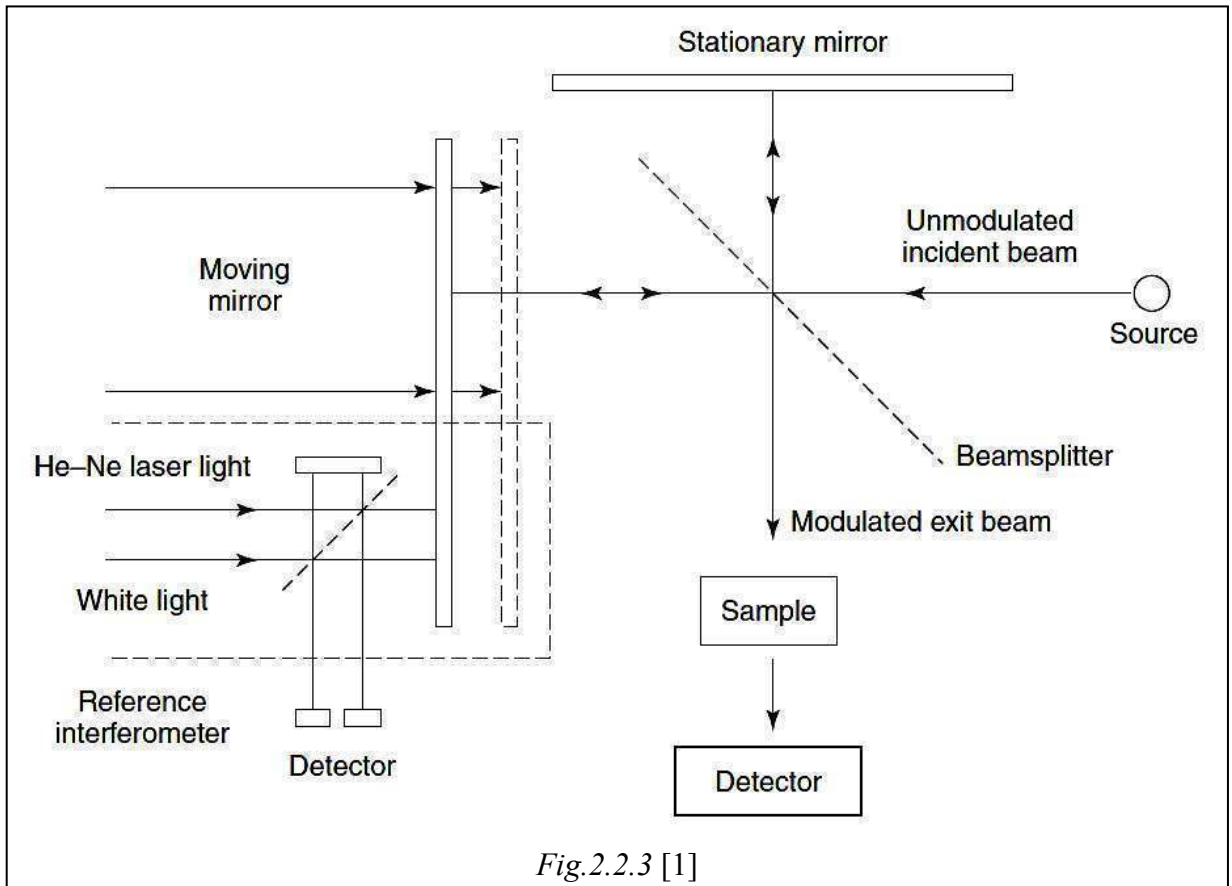


Fig.2.2.3 [1]

2.3.1 Beer-Lambert law

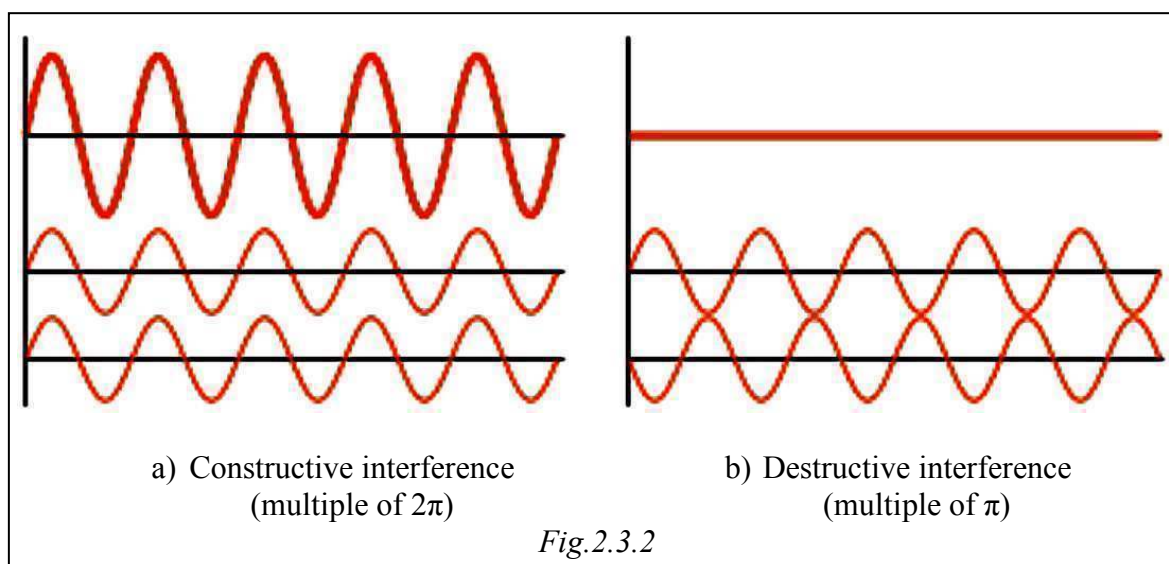
In infrared spectroscopy, infrared radiation is guided through the sample to be analyzed and the intensity of the transmitted radiation is measured as a function of wavelength. When the wavelength of light is matched to the energy difference between the equilibrium state and the excited state in a molecule of a sample, the light is absorbed. The absorption of light is given by the law of Beer-Lambert, whereby the absorbance A is given by the function:

$$A = c \cdot l \cdot \alpha(\nu)$$

where c is the concentration, I is the distance the light travels through the sample and $\alpha(\nu)$ is a specific molecular constant. From this function it is concluded that A increases linearly as a function of c and I . Generally, molecules vibrate in various ways thus resulting in the absorption of infrared light. This means that the constant $\alpha(\nu)$ depends on the wavelength. This explains how the concentrations of many different substances can be calculated simultaneously by the infrared spectroscopy. In the middle infrared spectroscopy, the fundamental molecular oscillations, which are usually narrow, are excited. Furthermore a substance displays different modes of oscillation at different frequencies thereby assisting the identification of atomic characteristics, even though there occurs some overlapping in frequencies. The study of the entire spectrum helps the distinction of chemically similar substances, such as ethanol and methanol as well as the quantitative determination of the two substances simultaneously.

2.3.2 Interferometer

The FTIR instruments use an interferometer for the measurement of the sample spectrum. This is the way the ERAspec model functions. A beam of light is emitted by a source. This beam is divided into two halves by a beam splitter consisting of ZnSe (zinc selenium). This allows the measurement of the spectrum in a range $1.5\text{-}18\ \mu\text{m}$ ($6660\text{-}550\ \text{cm}^{-1}$) which covers all the signals of interest for the analysis of fuels and lubricants. The two light beams generated are directed to a mirror reflecting the infrared light back towards the splitter where they recombine and directed to the detector. If both mirrors are located at the same distance from the splitter, no phase difference is observed. During a measurement one of the two mirrors is removed from the beam splitter with constant speed. When the relocation of the mirror becomes equal to a quarter of the light wavelength, the combined rays come out of phase, a fact that results in a destructive interference (*Fig.2.3.2*). In this case, no light impinges on the detector.



As the mirror moves, the light intensity measured by the detector will show an electronic configuration. In the field of electronics, configuration is the process of changing one or more characteristics of a periodic waveform of high frequency, defined as a carrier signal

based on another fluctuating signal. The intensity measured by the detector as a function of the mirror movement is the interferogram previously mentioned. For a monochromatic light source, the interferogram has a sinusoidal form. For multicolored light emitted from conventional infrared light sources, a representative value is calculated for a wide range of wavelengths. The interferogram will be the sum of the cosine waves of all frequencies present in the source. The Fourier transformation of the interferogram gives the spectrum. Matching the sample spectrum to that of the source of the sample, the absorption spectrum is generated. The absorption spectrum will show a range of frequencies that match the frequencies where the sample absorbed light. The spectrum analysis is a parameter which affects the clear separation of two adjacent frequency ranges in the final spectrum, which depends on how far the mirror moves. For a spectral analysis (i.e. analysis that will include the spectrum) of four wavenumbers (cm^{-1}) the mirror must be moved by 0.25 cm, whereas in order for the highest spectral resolution of 2 cm^{-1} to be achieved it must be moved by 0.5 cm.

2.3.3 Optimal operation of the Interferometer

The target for manufacturing an FTIR analyzer is always to measure the desirable samples with an optimal signal-to-noise ratio for more accurate quantitative analyses. This is called SNR (Signal to Noise Ratio) and is one of the basic principles of the instrument for repeatability reasons. The signal-to-noise ratio is analyzed into categories in the next subchapter. The careful selection and proper assembly of the spectrophotometer components, such as the light source, the detector, the interferometer and the electronics, are crucial for the of the construction of high quality instruments. Because absorption of light means its attenuation, very high absorption means that little light reaches the detector. This in turn suggests that too much absorption can not be measured accurately. On the other hand, if the absorption is very low, which means that little light is absorbed by the sample, the signal may be weaker compared to the noise of the instrument. In both cases the signal-to-noise ratio will limit the accuracy of the quantitative analysis. The optimum signal-to-noise ratio is when approximately 70% of the light is absorbed by the sample.

It is therefore evident from the formula $A = c \cdot l \cdot \alpha(\nu)$ that for molecules with very different $\alpha(\nu)$ or concentrations, the value of l (distance the light travels through the sample) of the system must be set so that the absorbance falls within a range where it can be measured with high accuracy. In order for adjacent frequencies in the spectra to be possible to distinguish, it is necessary to measure the spectrum at an appropriate resolution. *Fig.2.3.3.a* shows spectra of an ethanol-methanol mixture at different resolutions (2, 4, 8 and 16 cm^{-1}). If the resolution is too low the two signals (one for ethanol and one for methanol) can not be differentiated, i.e. to become distinct. It is impossible to distinguish them from one another at a spectral resolution of about 16 cm^{-1} . It can also be observed that the intensity of the signals decreases with reduced resolution. The spectrum analysis does not only affect the signal strength, but also the level of noise. *Fig.2.3.3.b* shows that the baseline noise increases by going from a spectral resolution of 4 to 2 cm^{-1} . For quantitative analysis the parameter that must be emphasized is the repeatability, i.e. small differences between consecutive measurements of the same sample. *Fig.2.3.3.c* shows the repeatability as a function of the resolution for MTBE (a gasoline additive). Higher resolutions lead to an increased deviation of the measured absorption impairing the repeatability of the measurement. These figures express strongly that there must be a golden section (compromise) between the repeatability of spectral separation of the overlapping frequencies.

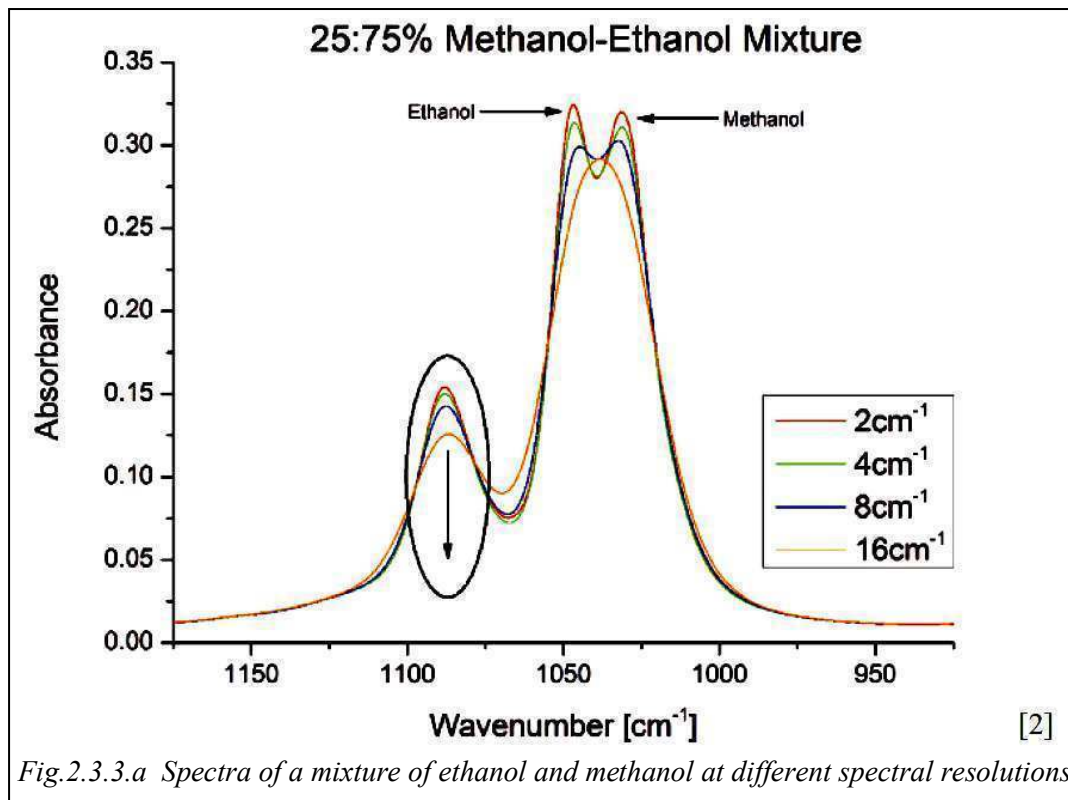


Fig.2.3.3.a Spectra of a mixture of ethanol and methanol at different spectral resolutions

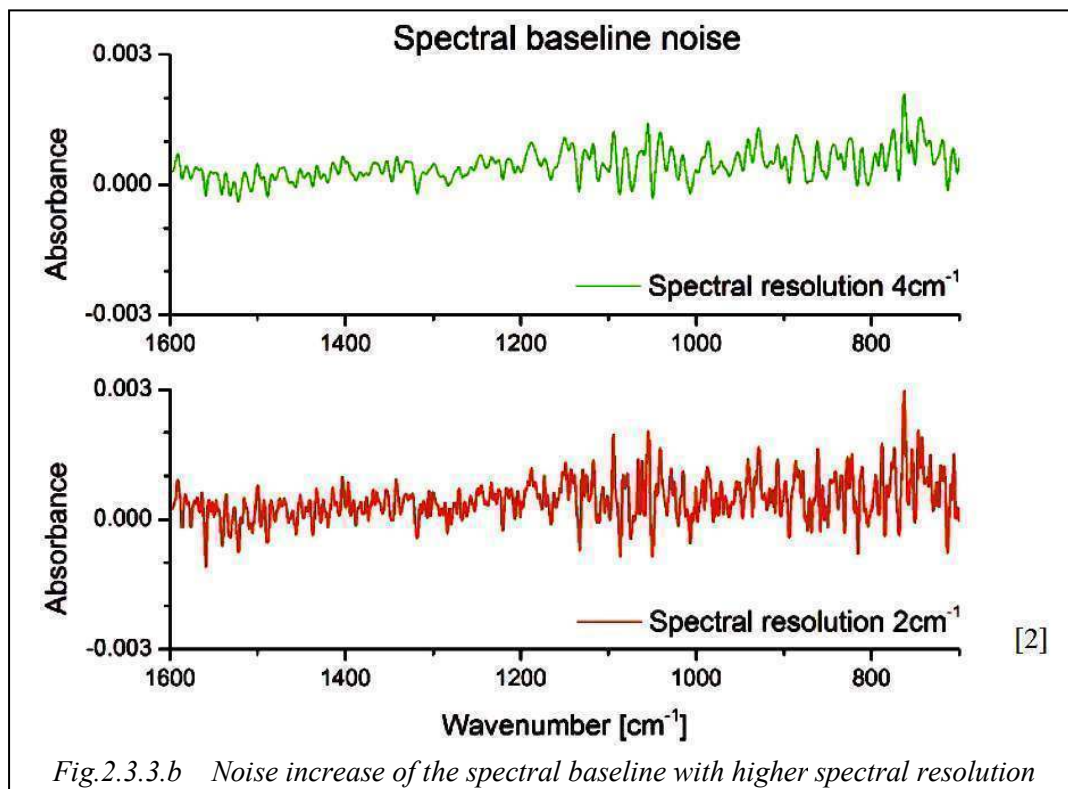


Fig.2.3.3.b Noise increase of the spectral baseline with higher spectral resolution

Observing Fig.2.3.3.d which shows a gasoline spectrum, it becomes obvious that no data of the spectrum is lost in spectral resolution 4 cm^{-1} . This is not surprising because the normal (linear) amplitude of vibrations for the molecules found in gasoline is larger than 10 cm^{-1} . For these frequency bands to become discrete, an analysis of the order of 4 cm^{-1} is sufficient,

as it is described in the ASTM standards. Increasing the resolution of the spectrum more, it just leads to reduction of the SNR as shown in *Fig.2.3.3.c*. Moreover, spectra in the spectral analysis of the order of 4 cm^{-1} are measured in half the time than at resolution of 2 cm^{-1} ,

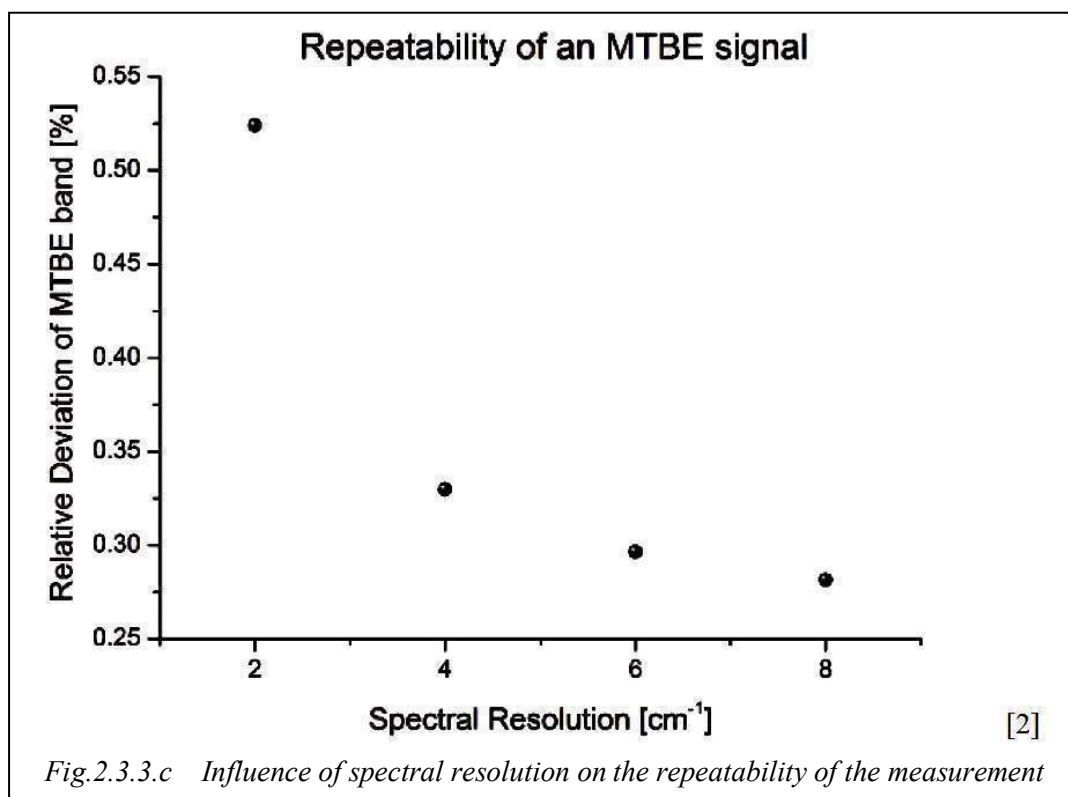


Fig.2.3.3.c Influence of spectral resolution on the repeatability of the measurement

because the moving mirror must cover shorter distance. This helps to shorten the analysis and increases the rate of process efficiency.

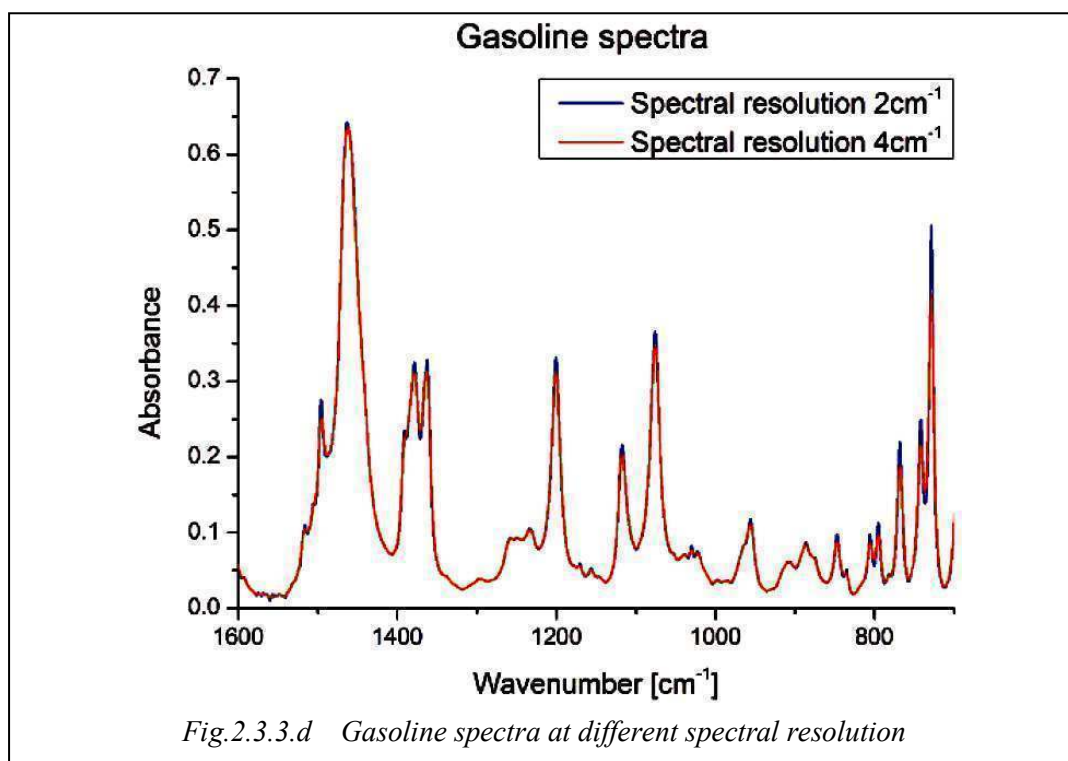


Fig.2.3.3.d Gasoline spectra at different spectral resolution

2.3.4 Mathematical proof of SNR

An FTIR spectrophotometer functions optimally when the sensor noise exceeds all the other sources of noise and is independent of the signal level. This is a common case for middle infrared spectroscopy but things may not be this way for shorter wavelengths. The sensitivity of the mid-infrared detectors is commonly expressed in terms of Noise Equivalent Power (NEP) of the detector, which is the ratio of the root mean square (rms) of the noise voltage V_n expressed in $V \cdot \text{Hz}^{1/2}$, to voltage response R_v of the detector expressed in $V \cdot W^{-1}$. It is an effective measure of the optical power given by a signal equivalent to the noise level. Therefore, the smaller the NEP, the more sensitive is the sensor. NEP is proportional to the square of the sensor surface area A_D (area of the photosensitive region of the detector), the constant of proportionality being known as the Specific Detectivity D^* :

$$D^* = \frac{A_D^{1/2}}{NEP} \quad \text{in } \text{cm} \cdot \text{Hz} \cdot \text{W}^{-1} \quad (\text{Equation 2.3.4.a})$$

(The unit $\text{cm} \cdot \text{Hz} \cdot \text{W}^{-1}$ sometimes is called Jones). The noise power N' which is observed in a measurement t (seconds) is given by:

$$N' = \frac{NEP}{T^{1/2}} \quad \text{in } W \quad (\text{Equation 2.3.4.b})$$

In order for the signal-to-noise ratio to be approached and also to be observable in a measurement, one must know not only the noise intensity, but also the signal intensity. The spectral brightness etc., the flow power per unit of area, per cubic angle in the wave number ν in a blackbody -source per temperature T - is given by the Planck equation:

$$U_\nu(T) = \frac{C_1 \cdot \nu^3}{\exp\left(C_2 \cdot \frac{\nu}{T}\right) - 1} \quad \text{in } W \cdot \text{sr}^{-1} \cdot \text{cm}^2 \cdot \text{cm}^{-1} \quad (\text{Equation 2.3.4.c})$$

where the C_1 and C_2 are the first and second angular constant, and have the values

$$C_1 = 2 \cdot h \cdot c^2 = 1.19110^{-12} \quad W \cdot \text{cm}^{-2} \cdot \text{sr} \cdot [(\text{cm}^{-1})^4]$$

$$C_2 = h \cdot \frac{c}{k} = 1.439 \text{K} \cdot \text{cm}$$

where k is the Boltzmann constant.

The intensity obtained by a detector through any optical system is determined by the "optical transfer" of the system and is the product of the area of the radius A , and its solid angle of

the decreasing diaphragm, $\Omega \cdot \text{sr}$. For a perfectly constructed FTIR spectrophotometer, the "optical transfer" is determined from the mirrors surface area of the interferometer and the maximum permitted stable angle, given by the equation:

$$\Omega_{\text{max}} = 2 \cdot \pi \cdot \alpha^2 = 2 \cdot \pi \cdot \frac{\Delta v}{v_{\text{max}}} \quad \text{in solid angles (steradians)} \quad (\text{Equation 2.3.4.d})$$

The intensity S' , emitted to a detector by an interferometer in a resolution $\Delta v \sim$ and operating with optical transfer Θ , with an effectiveness ξ to an extent of a wavenumber unit, is given by the equation:

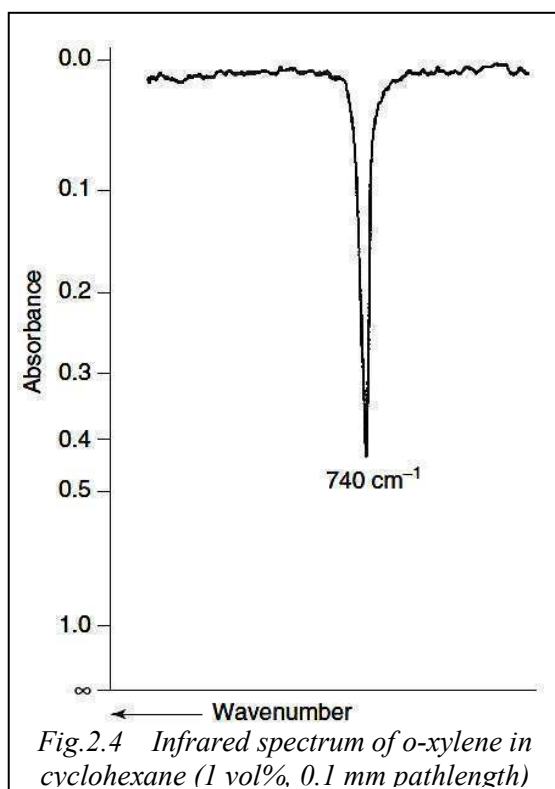
$$S' = U_v \cdot (T) \cdot \Theta \cdot \Delta v \cdot \xi \quad (\text{Equation 2.3.4.e})$$

Therefore the SNR of a spectrum measured with a dual beam interferometer is given by the equation:

$$\text{SNR} = \frac{S'}{N'} = \frac{U_v(T) \cdot \Theta \cdot \Delta v \cdot t^{0.5} \cdot \xi}{\text{NEP}} = \frac{U_v(T) \cdot \Theta \cdot \Delta v \cdot T^{0.5} \cdot D \cdot \xi}{(A_D)^{0.5}} \quad (\text{Equation 2.3.4.f})$$

This last equation (2.3.4.f) is the central relationship of signal-to-noise ratio. [12]

2.4 Analysis of general multicomponent mixtures



The analysis of a component within a complex mixture presents particular problems. In this subchapter there will be examined the approach of various components concentrations within a mixture. Below there are examples showing the fundamental ideas of the analysis of mixtures.

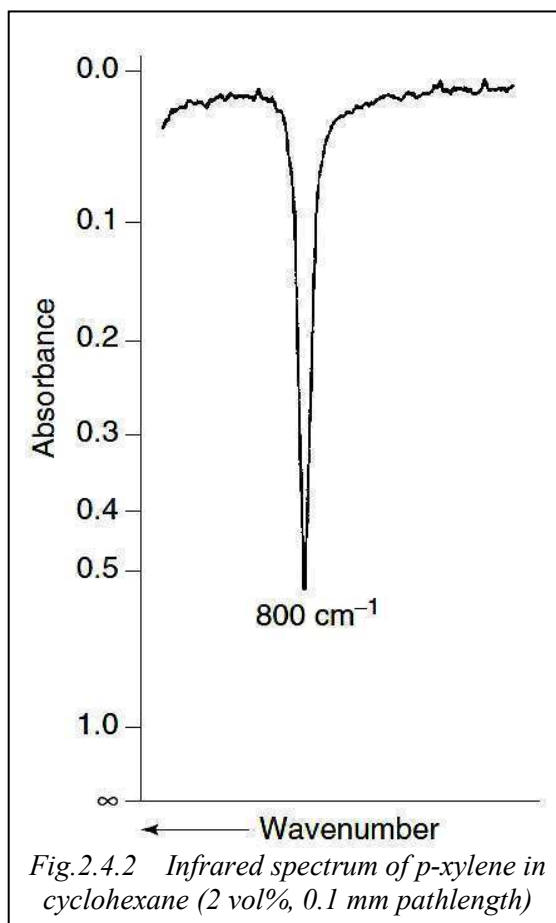
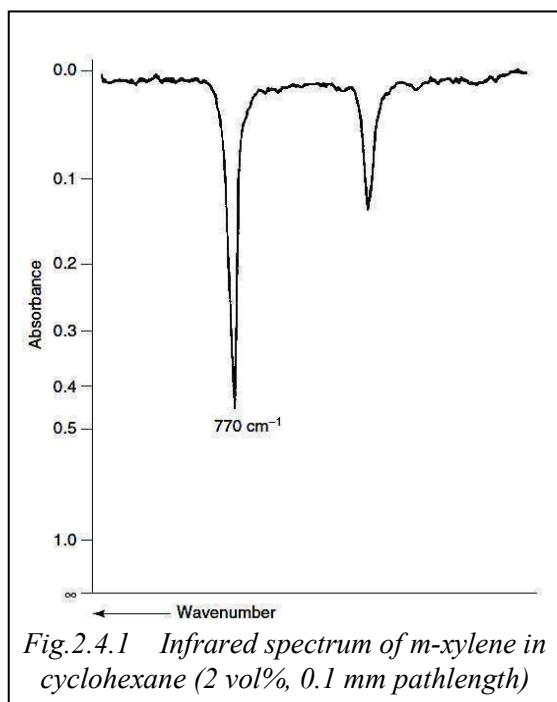
The example that follows, concerns the quantitative analysis of commercial xylene, which is a mixture of isomers, for example: 1,2-di-methyl-benzene (o-xylene), 1,3-di-methyl-benzene (m-xylene) and 1,4-di-methyl-benzene (p-xylene).

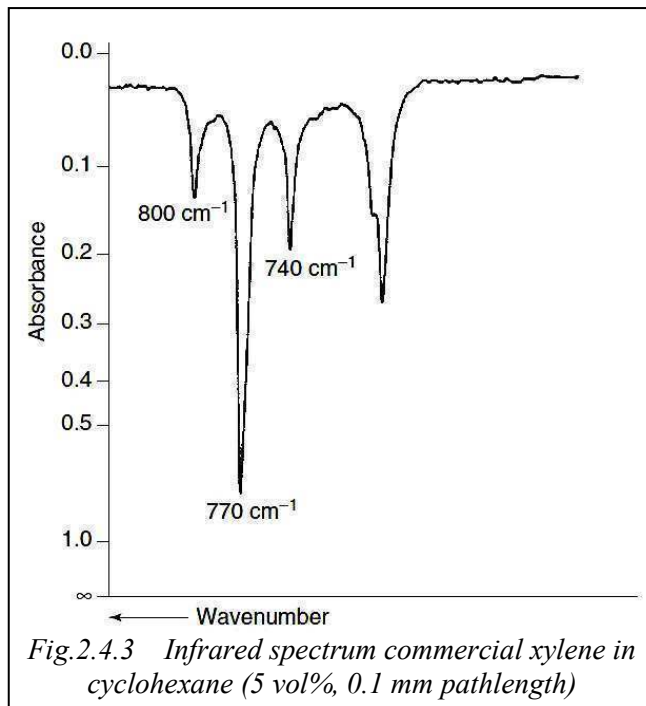
The spectra of these three pure xylenes dissolved in cyclohexane (Fig.2.4, 2.4.1, 2.4.2) show strong signals in the region 800-600 cm⁻¹. Cyclohexane has a very low absorption at this specific region, therefore being the proper solvent for the analysis. The infrared spectrum of commercial xylene is given in Fig.2.4.3.

The concentration of the three isomers should be determined in the sample of the market. First, the absorptions of xylenes at 740, 770 and 800 cm^{-1} from the standards shown in *Fig. 2.4, 2.4.1, and 2.4.2* must be measured as follows:

o-xylene	740	$A = 0.440 - 0.012 = 0.428$
m-xylene	770	$A = (0.460 - 0.015)/2 = 0.223$
p-xylene	800	$A = (0.545 - 0.015)/2 = 0.265$

The values for m-xylene and p-xylene were divided by two because their concentrations in the solution were double. Essentially the above calculations were based on the graphs. Absorbance values are also corrected for non-zero base in the spectral region.





The absorbance values are proportional to molecular absorption, and from now on the concentrations of xylenes can be determined in the mixture when the absorptions are measured. From Fig. 2.4.3 arise the following:

$$740 \quad A = 0.194 - 0.038 = 0.156$$

$$770 \quad A = 0.720 - 0.034 = 0.686$$

$$800 \quad A = 0.133 - 0.030 = 0.103$$

Dividing these absorbance values with the standard values previously exported, the content of each isomer in the mixture is taken by volume (vol%):

$$o\text{-xylene} = 0.156 / 0.428 = 0.364 \text{ vol\%}$$

$$m\text{-xylene} = 0.686 / 0.223 = 3.076 \text{ vol\%}$$

$$p\text{-xylene} = 0.103 / 0.265 = 0.389 \text{ vol\%}$$

It should be mentioned that there are potential errors in these values. There is an error due to the difficulty of starting line selection because of the overlapped peaks. In addition, every non-linearity of the Beer-Lambert law diagrams has been ignored, having used one value for the prediction of the constant for the proportionality with each of the above solutions. Alternatively, this same problem for xylene can be calculated by subtracting spectra. If a complex mixture is available, and also samples with its basic components, then the spectrum can be separated. This is achieved by removing a fraction i.e. the o-xylene, then the m-xylene and finally the p-xylene from the commercial xylene spectrum, so that the similar absorptions are zeroed.

Πίνακας 2.4

Case number	Values					
	Y	X ₁	X ₂	X ₃	X _p
1	y ₁	X ₁₁	X ₁₂	X ₁₃	X _{1p}
2	y ₂	X ₁₂	X ₂₂	X ₂₃	X _{2p}
.....
.....
n	y _n	X _{n1}	X _{n2}	X _{n3}	X _{np}

3. Methods of predicting properties / calibration

3.1 PCR method

It is obvious that the errors encountered in the previous process must not exist in measurements concerning important areas of production, such as medicines and fuel.

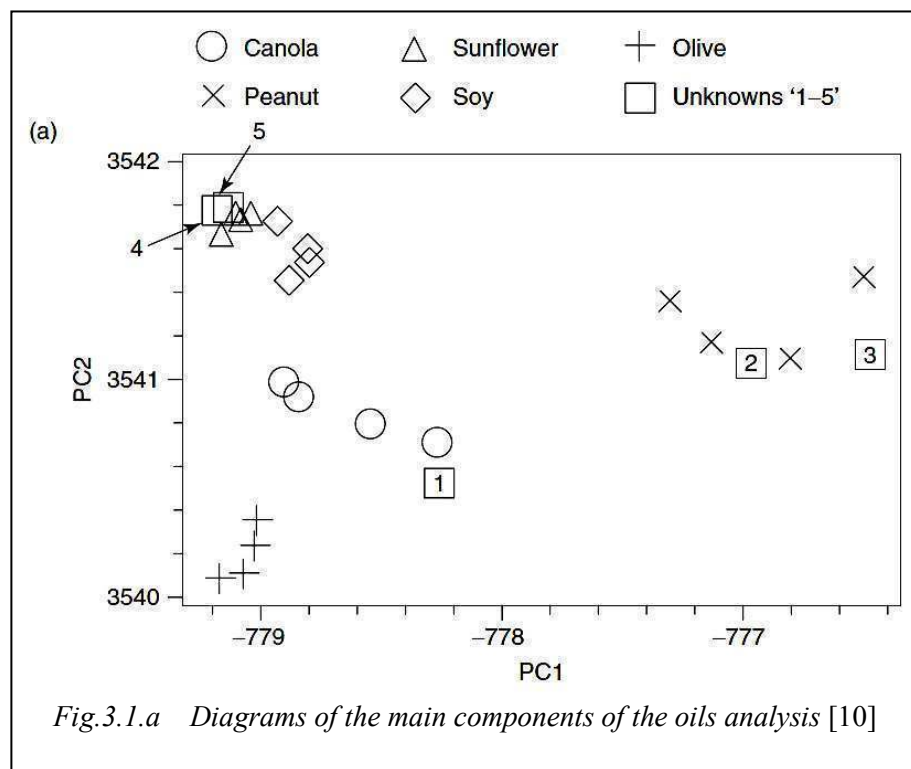
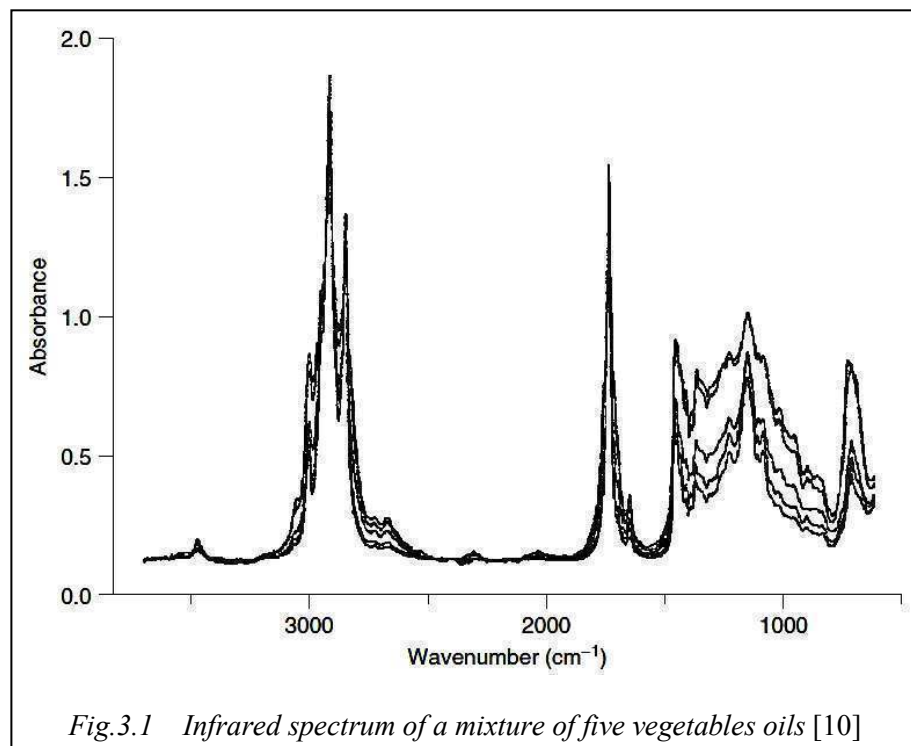
The improvements that have been made in computer technology associated with the spectroscopy have led to the expansion of the quantitative infrared spectroscopy. The applications of statistical methods in the analysis of experimental data are known as chemometrics ([3], [4], [5], [6], [7]). This object is very broad and there are too many bibliographies around it. In the following sections basic statistical predicting applications will be analyzed thoroughly, whereas at the same time three of them related directly to the instrument will be focused. These are: PCR, CA and MLR.

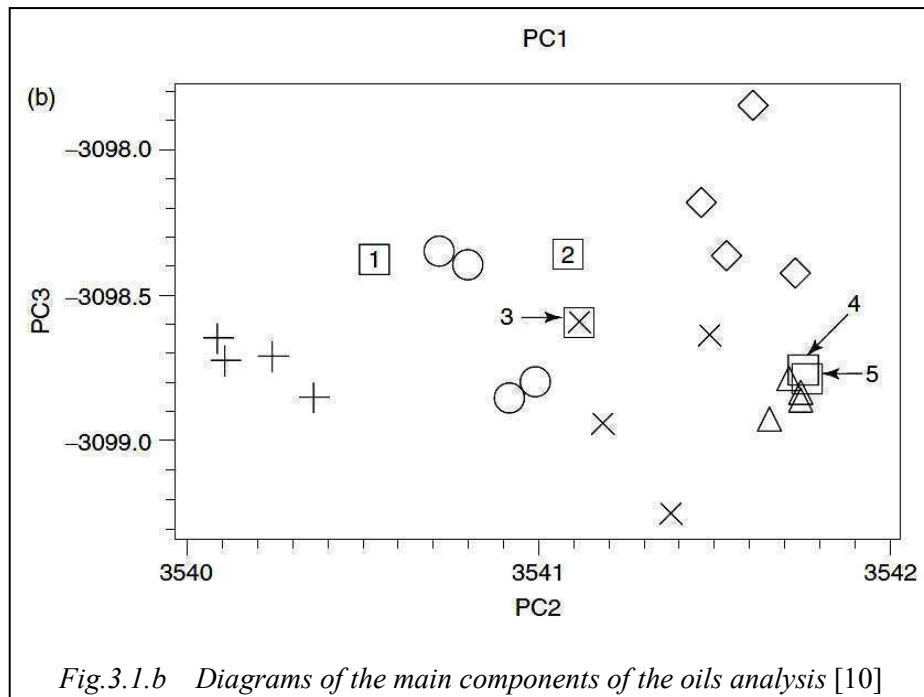
The most widely used in the analytical methods in infrared spectroscopy are: the classical of least squares (Classical Least Squares, CLS), the inverse least squares (Inverse Least Squares, ILS), the partial least squares (Partial Least Squares, PLS), the regression of the principal component (Principal Component Regression, PRC). CLS method (also known as K-matrix method) and PLS (known as the P-matrix method) are the least squares methods comprising calculation with tables. These methods are limited when complex mixtures are investigated. Methods functioning by factors, such as the PCR, are proven to be more useful.

Many times it is not necessary to know the exact concentrations of the components of interest, but simply whether these substances exist in the mixture or not. In these cases, a variable standard method of identification, such as the linear differential analysis (Linear Discriminant Analysis, LDA) or the Artificial Neural Networks (ANN's), may be used to identify characteristics of the spectra of the components of interest for the individual analyst [8]. Such methods are efficient to compare a large number of variables in a set of data, such as the intensity, the frequency and the amplitude. LDA and ANN's methods are known as supervisory methods, because there is some preexisting information for the sets (packets) of data. There are also non supervisory methods such as hierarchy by clusters (hierarchical clustering) which can be used to predict the data in a set of data without any preexisting information for them.

The use of calibration methods in infrared spectroscopy is easily understood generally from the application of PCR in vegetable oils [9]. *Fig.3.1* shows the infrared spectrum from 5 different types of vegetable oil, added one over the other. This demonstrates the similarity of these spectra and the difficulty in finding differences between the samples by simple observation. However, there are marked differences in the range of the positions of each spectrum which someone can take advantage of by using the PCR. The wavenumbers of the C-H bond when stretched and when bent in the fingerprint area can be used as constants. The analysis process includes a series of steps. First you need to get the spectrum for the oils and for the unknown samples. The spectra must be identical, i.e. they have to be taken under the same conditions. Then the average of each wavenumber for each category of oil is calculated. Immediately after that, the absolute difference between the unknown wave number and the average of each class and each "line group" (band) is calculated. One step

further, the sum of the differences is calculated for each class between all line groups, and the identification of the unknown with the class of the smallest sum of differences, can be made. The results of the analysis of a mixture of unknown vegetable oils, using seven peaks, are shown in *Fig.3.1*.





Different symbols were used to represent each type of known oil (eg. ○ for Canola) and the small square (□) for the unknowns. The diagram of Fig.3.1.a shows a PC1 against a PC2 diagram and shows that canola oil and peanut oil form distinct clusters. This graph can be used to determine the unknown "1" with the type of canola oil, and the unknowns "2" and "3" with the type of peanut oil. The graph shown in Fig.3.1.b shows a PC3 against a PC2 diagram. In this case sunflower oil and soybean oil are clearly separated while in the previous comparison (PC2 versus a PC1 graph) were not clear. The unknowns "4" and "5" may be determined as sunflower oil because they form a cluster near the known samples of sunflower oil in both diagrams.

3.2 Cluster Analysis method

The terms cluster, group and class, are used in a completely instinctive way in bibliographies without significant effort of interpretation. In essence it concluded that the official interpretation of the term is not only difficult but sometimes it is interpreted incorrectly. For example, in 1964 Bonner suggested that the ultimate criterion for assessing the meaning of such notions is the user's judgment. If someone using a term such as "cluster" gives the researcher an answer of worth, this is all needed.

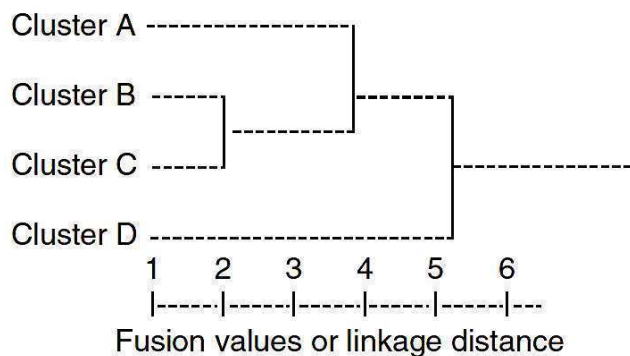
Bonne's view is partly acceptable but not convincing to many writers like Cormak (1971) and Gordon (1999) who tried to interpret what "cluster" means in terms of internal consistency, homogeneity and external isolation-separation.

3.2.1 Simple presentation of a cluster analysis

The classification with hierarchy is perhaps the best known statistical method for finding relatively homogeneous clusters in cases based on measured characteristics. It starts with

each case as a separate cluster, i.e.: there are so many clusters as are the cases, and then combines the clusters continually by reducing their number in each step until only one cluster remains. This method uses the dissimilarities or the distances between objects (objects under classification) when creating clusters. The SPSS program calculates "distances" between the points of data in terms of separate variables.

A hierarchical tree diagram, called a dendrogram in SPSS, can be drawn to show the points that become links. The clusters are connected in increasing serial levels of dissimilarity. The measurement of dissimilarity itself is dependent on the measure used, for example:



This example shows clusters B and C to be combined at the fused value of 2 and B and C with A, at 4. The fused values or the distances of the links are calculated with SPSS. The purpose of the algorithm in clusters is to join objects together successively in larger clusters, using some measure of similarity or distance. To the left of dendrogram begins with each item or case in a class by itself (in our example there are only four cases). With very small steps, the user makes the criterion "mild" of what is and is not unique. In other words the user considers increasing differences acceptable in order to decide if two or more items will be members of the same cluster.

As a result more and more items are connected together and larger clusters fuse with increasingly dissimilar items within them. In the end, at the last step, all the items are linked together as one cluster. In these graphs, the horizontal axis indicates the fusion or the distance of the linkage. For each node in the graph (where a cluster is created) the user can read the distance criterion with which similar items were linked together in a new cluster. As a general procedure, "clustering" can be summarized as follows:

- The distance is calculated between all of the original clusters. In most analyses, the initial clusters will contain only one item.
- Then, the two more matching clusters are fused and the distances are redefined.
- The previous step is repeated until all the items result in a cluster.

This study is very general but is the closest to that of the instrument. The corresponding distance at Eraspec data is the Mahalanobis distance. These will be analyzed in a following chapter where the cluster analysis fully customized as it happens to the instrument, will be explained.

3.3 Multi Linear Regression (MLR) method

In multiple regression, many independent variables are used for making a model of a response variable. For each one of the n cases that have been studied, there are values collected for the response and for each one of the independent variables. If the response is Y , and the independent variables are X_1, X_2, \dots, X_p (p is the number of the independent variables, or of those that will "make the prediction", then the data will form the: $n \times (p+1)$ table:

For simple regression, $p=1$. In the presentation of data, the value x_{ij} refers to the value for the j -th variable in the i -th case. The values for a case are shown in a row, while all values for a variable in a column.

In multiple regression, an equation which expresses the response as a linear function of p independent variables is estimated using the observed data. The model is determined by a linear equation,

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + e \quad (\text{Equation 3.3})$$

where β are the unknown parameters, e are the statistical errors, Y is the dependent variable or the response, and X_1, X_2, \dots, X_p are those that make the prediction. When $p=2$, the equation 3.3 gives a plane in a three dimensional space (X_1, X_2, Y).

The MLR will be studied in a following chapter adapted to the needs of the instrument, as the user can intervene and change the model of the method. [12]

4. Application of FTIR in organic molecules

4.1 Introduction

One of the most common applications of infrared spectroscopy is the identification of organic compounds. Besides, this is also what the instrument under calibration does. Following are studied the major categories of substances found in gasoline and diesel, but also others being of industrial interest since the organic compounds are widely used as raw materials.

4.2 Aliphatic hydrocarbons

Aliphatic hydrocarbons, alkanes, contain only C-H and C-C bonds, but much information can be gained from the infrared spectra that these molecules give. The most useful are those resulting from the C-H bonds when they stretch and also when bent. The regions where C-H bonds bend in aliphatic hydrocarbons are $3000\text{-}2800\text{ cm}^{-1}$. C-H bonds that stretch in methyl

and methylene groups are diversified. For methyl groups, the asymmetric C-H bond absorbs at 2870 cm^{-1} while the corresponding symmetrical at 2960 cm^{-1} . By comparison, methylene groups absorb in the asymmetric stretching at 2930 cm^{-1} and in the symmetric stretching at 2850 cm^{-1} . The bending C-H bond increases its activity at the range below 1500 cm^{-1} . Methyl groups produce "two ranges" where bending appears, i.e. a symmetrical at 1380 cm^{-1} and an asymmetrical at 1475 cm^{-1} . Methylene groups on the other hand absorb in four types of vibrations: "Scissoring" at 1465 cm^{-1} , "Rocking" at 720 cm^{-1} , "Wagging" at 1305 cm^{-1} and "Twisting" at 1300 cm^{-1} . The intensity of methylene CH_2 when in vibration of the "rocking" type is useful, since four or more CH_2 groups are necessary in a chain, in order to produce a distinct peak near 720 cm^{-1} . Shorter chains show a more variable peak, for example for CH_2 the "rocking" peak in C_4H_{10} is near 734 cm^{-1} . Although these are the main absorptions associated with the aliphatic hydrocarbons, there is a large number of peaks appearing in the spectrum of such components because there is also a wide range of possible compounds (of different structure). The main infrared absorbances for the alkanes are summarized in the following table:

Wavenumber (cm^{-1})	Assignment
<i>Alkanes</i>	
2960	Methyl symmetric C-H stretching
2930	Methylene asymmetric C-H stretching
2870	Methyl asymmetric C-H stretching
2850	Methylene symmetric C-H stretching
1470	Methyl asymmetrical C-H bending
1465	Methylene scissoring
1380	Methyl symmetrical C-H bending
1305	Methylene wagging
1300	Methylene twisting
720	Methylene rocking
<i>Alkenes</i>	
3100–3000	=C-H stretching
1680–1600	C=C stretching
1400	=C-H in-plane bending
1000–600	=C-H out-of-plane bending
<i>Alkynes</i>	
3300–3250	=C-H stretching
2260–2100	C≡C stretching
700–600	=C-H bending

Alkenes have the double bond group, the majority of which also contain hydrogen in the double bond carbon. Four large molecular disorders are associated with these pieces. These

are the inside and outside level =C-H conversions, C=C stretching and =C-H stretching. The characteristic absorptions observed in alkenes are also seen in *Table 4.2*.

Alkynes contain the C≡C group and three characteristic bands can be present, i.e. ≡C-H stretching, ≡C-H bending and C≡C stretching. The main absorbances observed for alkynes are referred in *Table 4.2*.

4.3 Aromatic components

The aromatic components show characteristic peaks useful in five regions of the medium infrared spectrum. These are shown in *Table 4.3* below:

Wavenumber (cm ⁻¹)	Assignment
3100–3000	C–H stretching
2000–1700	Overtone and combination bands
1600–1430	C=C stretching
1275–1000	In-plane C–H bending
900–690	Out-of-plane C–H bending

The C-H bond stretching in aromatic components is displayed in a range between: 3100-3000 cm⁻¹. This renders them easy to differentiate from those derived from aliphatic C-H groups which appear below 3000 cm⁻¹. In the region 2000-1700 cm⁻¹ a series of weak combinations and harmonic

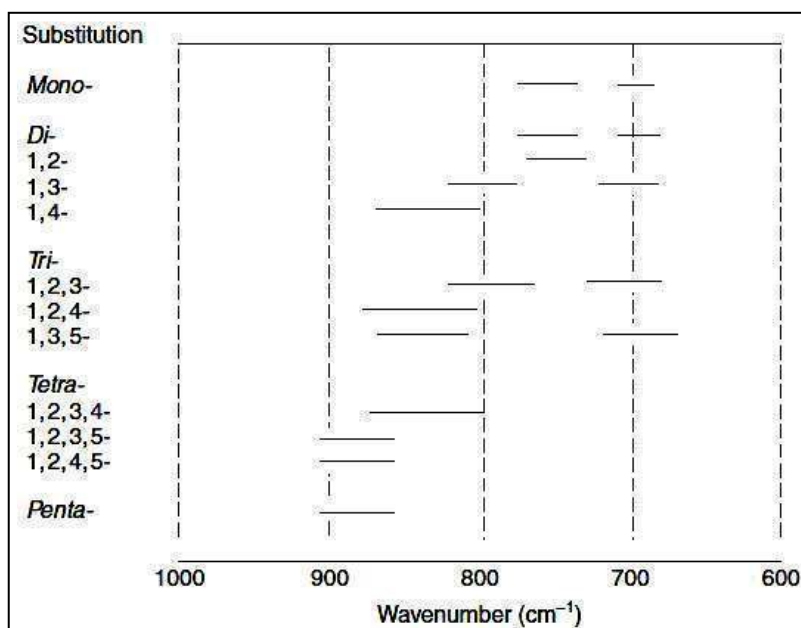


Fig. 4.3.a Absorption due to change in the angle of the aromatic molecule

groups of peaks appear indicating the existence of benzene. Skeletal vibrations representing the stretch of the C=C bond absorb in a range between 1275-1000 cm⁻¹ (in plane bending) and 900-690 cm⁻¹ (out of plane bending). The vibrations of the bending out of plane of the aromatic components are strong and characteristic of the number of hydrogens in the ring, and therefore can be used to replace the standard. This information is summarized in the table of *Fig. 4.3.a*.

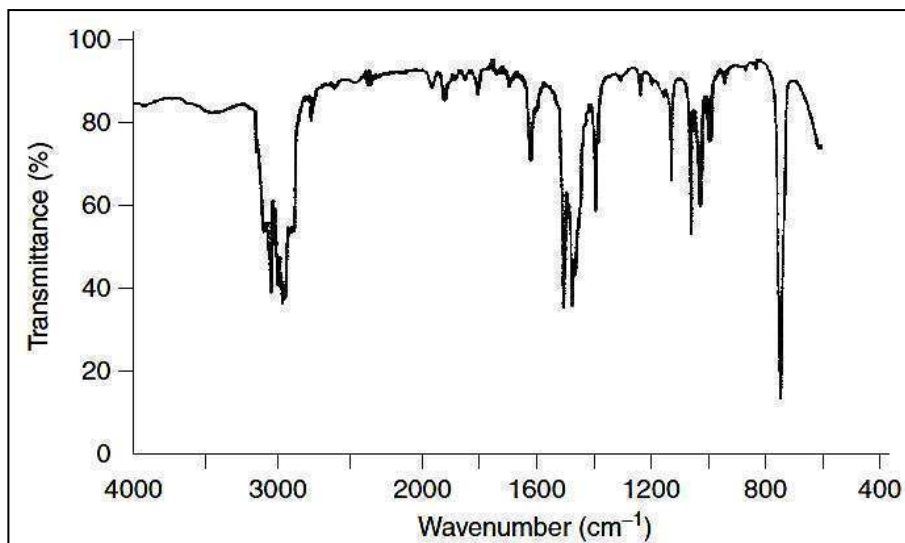


Fig.4.3.b Infrared spectrum of a substituted benzene molecule: element A

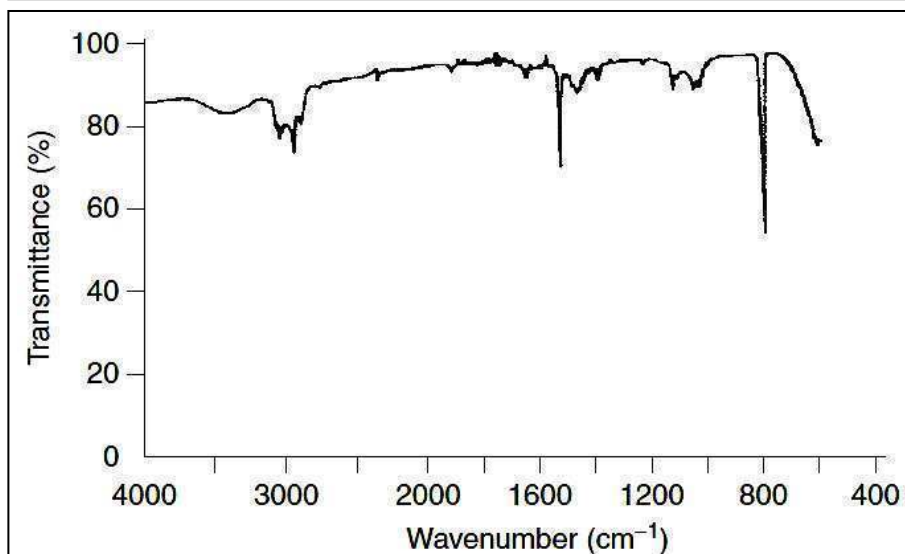


Fig.4.3.c Infrared spectrum of a substituted benzene molecule: element B

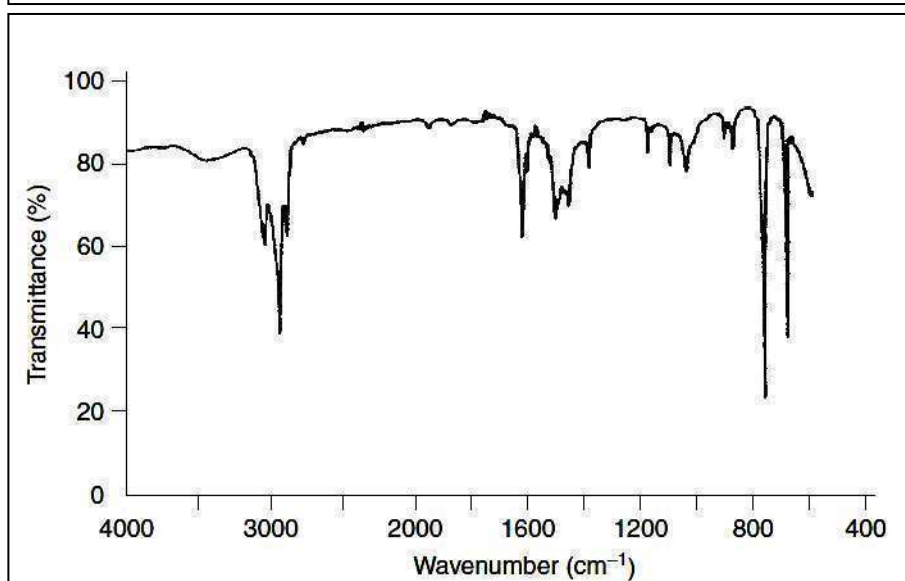


Fig.4.3.d Infrared spectrum of a substituted benzene molecule: element C

4.4 Oxygen containing components

4.4.1 Alcohols and phenols

Alcohols and phenols produce characteristic peaks in the infrared due to O-H and C-O stretching bonds, which both are sensitive to hydrogen bonds. For alcohols, the wide O-H bond lies in the range of 3600 cm^{-1} , while in phenols this range appears at $50\text{-}100\text{ cm}^{-1}$, lower than for alcohols. The stretching of C-O bond in the alcohols and phenols produces a strong signal in the region $1300\text{-}1000\text{ cm}^{-1}$. These compounds also produce O-H vibrations due to stretching, but the second "mates" with other vibrations and produces strange peaks in the region of the fingerprint. The main absorptions for these items are presented in *Table 4.4.1* below.

Wavenumber (cm^{-1})	Assignment
	<i>Alcohol and phenols</i>
3600	Alcohol O-H stretching
3550–3500	Phenol O-H stretching
1300–1000	C–O stretching
	<i>Ethers</i>
1100	C–O–C stretching
	<i>Aldehydes and ketones</i>
2900–2700	Aldehyde C–H stretching
1740–1720	Aliphatic aldehyde C=O stretching
1730–1700	Aliphatic ketone C=O stretching
1720–1680	Aromatic aldehyde C=O stretching
1700–1680	Aromatic ketone C=O stretching
	<i>Esters</i>
1750–1730	Aliphatic C=O stretching
1730–1705	Aromatic C=O stretching
1310–1250	Aromatic C–O stretching
1300–1100	Aliphatic C–O stretching
	<i>Carboxylic acids</i>
3300–2500	O–H stretching
1700	C=O stretching
1430	C–O–H in-plane bending
1240	C–O stretching
930	C–O–H out-of-plane bending
	<i>Anhydrides</i>
1840–1800	C=O stretching
1780–1740	C=O stretching
1300–1100	C–O stretching

4.4.2 Ethers

Ethers can be detected by a strong C-O signal due to stretching of the bond near 1100 cm^{-1} , because of the C-O-C bond in this kind of bonding (*Table 4.4.1*). Aromatic ethers give a strong signal near 1250 cm^{-1} while cyclic ethers at $1250\text{-}900\text{ cm}^{-1}$.

4.4.3 Aldehydes and ketones

Aliphatic and aromatic ketones show carbonyl absorptions at $1730\text{-}1700$ and $1700\text{-}1680\text{ cm}^{-1}$ respectively, while the aliphatic and aromatic aldehydes produce carbonyl absorptions at $1740\text{-}1720\text{ cm}^{-1}$ respectively (*Table 4.4.1*). The position of the C=O wavenumber ranging within these limits depends on the bonds with hydrogens and the coupling with them in the molecule. Coupling with a C=C bond leads to the change in position of the group C=O, which causes the absorption to fall at lower wavenumber. Aldehydes also show a characteristic C-H stretching bond ranging at $2900\text{-}2700\text{ cm}^{-1}$.

4.4.4 Esters

The two most polar bonds in esters included in (-CO-OC-) is C=O and C-O, and these bonds produce the strongest lines in the spectrum of each ester (summarized in *Table 4.4.1*). Aromatic and aliphatic esters may be differentiated, since both C=O stretching and C-O stretching produce lines at $1750\text{-}1730\text{ cm}^{-1}$ and $1300\text{-}1100\text{ cm}^{-1}$ respectively, while aromatic esters produce lines in the spectrum due to C=O and C-O at $1730\text{-}1705\text{ cm}^{-1}$ and $1310\text{-}1250\text{ cm}^{-1}$ respectively.

4.5 Tables of various wavenumbers

In this subchapter presented are tables displaying the essence and the value of the wavenumber where the absorption takes place in secondary compounds.

4.5.1 Amines

Wavenumber (cm^{-1})	Assignment
3335	N-H stretching (doublet for primary amines; singlet for secondary amines)
2780	N-CH ₂ stretching
1615	NH ₂ scissoring, N-H bending
1360–1250	Aromatic C-N stretching
1220–1020	Aliphatic C-N stretching
850–750	NH ₂ wagging and twisting
715	N-H wagging

4.5.2 Amides

Table 4.5.2
Characteristic wavenumbers of amides

Wavenumber (cm ⁻¹)	Assignment
2260–2240	Aliphatic nitrile C≡N stretching
2240–2220	Aromatic nitrile C≡N stretching
2180–2110	Aliphatic isonitrile –N≡C stretching
2160–2120	Azide N≡N stretching
2130–2100	Aromatic isonitrile –N≡C stretching
1690–1620	Oxime C=N–OH stretching
1680–1650	Nitrite N=O stretching
1660–1620	Nitrate NO ₂ asymmetric stretching
1615–1565	Pyridine C=N stretching, C=C stretching
1560–1530	Aliphatic nitro compound NO ₂ asymmetric stretching
1540–1500	Aromatic nitro compound NO ₂ asymmetric stretching
1450–1400	Azo compound N=N stretching
1390–1370	Aliphatic nitro compound NO ₂ symmetric stretching
1370–1330	Aromatic nitro compound NO ₂ symmetric stretching
1300–1270	Nitrate NO ₂ symmetric stretching
965–930	Oxime N–O stretching
870–840	Nitrate N–O stretching
710–690	Nitrate NO ₂ bending

4.5.2.a Compounds with nitrates

Table 4.5.2.a
Characteristic wavenumbers of compounds with nitrates

Wavenumber (cm ⁻¹)	Assignment
3360–3340	Primary amide NH ₂ asymmetric stretching
3300–3250	Secondary amide N–H stretching
3190–3170	Primary amide NH ₂ symmetric stretching
3100–3060	Secondary amide amide II overtone
1680–1660	Primary amide C=O stretching
1680–1640	Secondary amide C=O stretching
1650–1620	Primary amide NH ₂ bending
1560–1530	Secondary amide N–H bending, C–N stretching
750–650	Secondary amide N–H wagging

4.5.3 Compounds containing halogens

Wavenumber (cm ⁻¹)	Assignment
1300–1000	C–F stretching
800–400	C–X stretching (X = F, Cl, Br or I)

4.5.4 Compounds containing boron

Wavenumber (cm ⁻¹)	Assignment
3300–3200	B–O–H stretching
2650–2350	B–H stretching
1465–1330	B–N stretching
1380–1310	B–O stretching
1205–1140	B–H bending
980–920	B–H wagging

4.5.5 Compounds containing phosphorus

Wavenumber (cm ⁻¹)	Assignment
2425–2325	Phosphorus acid and ester P–H stretching
2320–2270	Phosphine P–H stretching
1090–1080	Phosphine PH ₂ bending
990–910	Phosphine P–H wagging
2700–2100	Phosphorus acid and ester O–H stretching
1040–930	Phosphorus ester P–OH stretching
1050–950	Aliphatic asymmetric P–O–C stretching
830–750	Aliphatic symmetric P–O–C stretching
1250–1160	Aromatic P–O stretching
1050–870	Aromatic P–O stretching
1450–1430	Aromatic P–C stretching
1260–1240	Aliphatic P=O stretching
1350–1300	Aromatic P=O stretching
1050–700	P–F stretching
850–500	P=S stretching
600–300	P–Cl stretching
500–200	P–Br stretching
500–200	P–S stretching

4.5.6 Compounds containing sulphur

Wavenumber (cm ⁻¹)	Assignment
700–600	C–S stretching
550–450	S–S stretching
2500	S–H stretching
1390–1290	SO ₂ asymmetric stretching
1190–1120	SO ₂ symmetric stretching
1060–1020	S=O stretching

5. Fuel analysis using FTIR

Fuels are complex mixtures of alkanes (paraffins), alkenes (olefins) and aromatics. Additionally, it is enriched with a wide range of additives to improve combustion and reduce emissions parameters. Therefore substances enriching it with oxygen are usually used, such as MTBE, methanol or ethanol, which are easily checked with infrared spectroscopy according to ASTM D5845 standard. Other ingredients must also be checked because this is strictly defined by international rules since environmental and health problems are concerned.

Benzene for example can be monitored in accordance with ASTM D6277, EN 238 standards using infrared spectroscopy. In recent years the growing need for the use of renewable fuel sources led to mixing of bioethanol in gasoline. FAME an ester derived from plant oil, often is added in diesel as a deactivator of carbon dioxide (ASTM D7806, EN 14078). All these components exist in modern fuel and usually at high concentrations. However there are fuel additives such as MMT or 2-EHN which have intense effect on the properties of the fuel already at very low concentrations. For example, 50 mg/L of MMT increase the octane number (RON) usually by 2.5 units and 250 mg/L of 2-EHN increase the cetane number by 1.7 units. Therefore it is very important to measure these parameters with the utmost precision, along with other additives which are present in much higher concentrations.

As described earlier, longer distance traveled by the transmission cell leads to higher signals, which allow very low detection limits. Ingredients of high concentrations lead to intense or complete absorption of light if the traveled distance is too long. ERASPEC, in order to be able to measure additives accurately which are either in low or in high concentrations, has a double system of cells. Normally, the additives present in high concentrations such as ethanol (up to 25 Vol%) are measured in a 20 μm cell while the additives present in small concentrations such as MMT (added to ppm range) are measured in a 100 μm cell. The concentrations of components though, are not the only properties of interest in fuel analyzes. There also exist other parameters, such as RON, MON, the distillation properties, bubble

point and vapor pressures, which are of great interest as well. These fuel parameters can not match specific signals in the spectrum because they do not match a single ingredient only. Using many different signals and applying advanced statistical methods such parameters can be estimated from a spectrum. For more correct results it is not enough only to measure the spectrum as accurately as possible, but also the instrument must use spectra database for the matching of the parameters. Therefore, the result also depends on the quality of the reference data, for example, the anti-knocking effect of the engine for octane numbers or the vapor pressure, are measured by means of one of the available methods. To obtain reliable "predictions" using such a model, the predicted value should "fall" in the calibrated range of the model. If for calibration only samples with an octane number between 90 and 92 have been used, the device will not be able to export correct results for gasoline with an octane number of 95. In addition, given the complexity of fuels and the diversity they appear from area to area, is not possible to develop a universal database. The ERASPEC device has been installed in many regions of the world and therefore it offers widely enriched databases for different regions in order to achieve more accurate results. In the next two subchapters the properties of gasoline and oil are analyzed with regards to the device.

5.1 Gasoline analysis

The ERASPEC strictly complies in all aspects with the requirements of the latest American and European standards ASTM D 5845, ASTM D 6277 and EN 238 for the measurement of the following parameters:

15 Aromatics:

Benzane (ASTM D 6277, EN 238)	3-Ethyl toluene
Toluene	4-Ethyl Toluene
o-Xylene	Mesitylene
p-Xylene	Pseudocumene
m-Xylene	Hemellitol
Ethyl benzene	Durene
Propyl benzene	Napthalene
2-Ethyl toluene	

11 Oxygenates (ASTM D 5845):

4 Ethers:

MTBE
TAME
ETBE
DIPE

5 Alcohols

Methanol
Ethanol
Iso-Propanol
2-Butanol
tert-Butanol

Other Oxygenates

Dimethoxymethane (DMM)
Dimethylcarbonate (DMC)

Total Parameters:

Oxygen	Saturates
Oxygenates	Diolefins

Aromatics
Olefins

Anilines

Other Parameters:

MMT, CMT, & Manganese
DCPD (Dicyclopentadiene)

Aniline
N-Methylaniline

The density is measured from 0.500 to 1.999 (direct measurement with an internal densitometer of the Anton Paar firm according to IP 599, correl. (ASTM D 4052)). The properties are: Research Octane Number (RON-correl. ASTM D 2699), Motor Octane Number (MON-correl. ASTM D 2700), Anti Knock Index (AKI), Distillation points (correl. ASTM D86): IBP, T10, T50, T90, FBP. Evaporation points (correl. ASTM D86): E200, E300. RVP (correl. ASTM D 5191), Driveability Index (correl. ASTM D4814). VOC Emissions Calculator, and ten more additional freely programmable parameters.

The above in tabular form with the traceability and repeatability limits recorded, are summarized as follows: *Table 5.1.a* and *Table 5.1.b* [13]

Table 5.1.a

Substance	Range	Repeatability	at:	Detection Limit
Benzene	0 – 10 Vol%	0.018 Vol%	@ 0.5 Vol%	0.06 Vol%
		0.024 Vol%	@ 1 Vol%	
		0.07 Vol%	@ 2 Vol%	
		0.1 Vol%	@ 5 Vol%	
Toluene	0 – 20 Vol%	0.3 Vol%	@ 7 Vol%	0.5 Vol%
m-Xylene	0 – 20 Vol%	0.2 Vol%	@ 5 Vol%	0.6 Vol%
o-Xylene	0 – 20 Vol%	0.2 Vol%	@ 6 Vol%	0.6 Vol%
p-Xylene	0 – 20 Vol%	0.2 Vol%	@ 6 Vol%	0.7 Vol%
Ethylbenzene	0 – 20 Vol%	0.2 Vol%	@ 6 Vol%	0.7 Vol%
Propylbenzene	0 – 20 Vol%	0.2 Vol%	@ 5 Vol%	0.8 Vol%
4-Ethyltoluene	0 – 20 Vol%	0.2 Vol%	@ 5 Vol%	0.6 Vol%
Pseudocumol	0 – 20 Vol%	0.2 Vol%	@ 5 Vol%	0.7 Vol%
Mesitylene	0 – 20 Vol%	0.2 Vol%	@ 5 Vol%	0.6 Vol%
All other Aromatics	0 – 20 Vol%	0.2 Vol%	@ 5 Vol%	0.7 Vol%
MTBE	0 – 20 Vol%	0.08 Vol%	@ 6 Vol%	0.5 Vol%
		0.2 Vol%	@ 13 Vol%	
ETBE	0 – 25 Vol%	0.1 Vol%	@ 9 Vol%	0.5 Vol%
TAME	0 – 25 Vol%	0.1 Vol%	@ 11 Vol%	0.5 Vol%
DIPE	0 – 20 Vol%	0.1 Vol%	@ 10 Vol%	0.5 Vol%
Methanol	0 – 15 Vol%	0.05 Vol%	@ 5 Vol%	0.5 Vol%
Ethanol	0 – 25 Vol%	0.15 Vol%	@ 10 Vol%	0.5 Vol%
		0.2 Vol%	@ 24 Vol%	
tert-Butanol	0 – 25 Vol%	0.1 Vol%	@ 11 Vol%	0.5 Vol%
All other Oxygenates	0 – 25 Vol%	0.1 Vol%	@ 10 Vol%	0.5 Vol%
Total Oxygen	0 – 12 Vol%	0.04 Vol%	@ 1.6 Vol%	0.09 Vol%

Table 5.1.a (cont.)

Total Olefins	0 – 80 Vol%	0.7 Vol%	@ 15 Vol%	0.5 Vol%
Total Aromatics	0 – 50 Vol%	0.8 Vol%	@ 33 Vol%	0.5 Vol%
Saturates	0 – 100 Vol%	1.1 Vol%	@ 45 Vol%	Not applicable
Di-Olefins	0 – 15 Vol%	0.18 Vol%	@ 5 Vol%	0.2 Vol%
DCPD (Dicyclopentadiene)	0 – 15 Vol%	0.11 Vol%	@ 5 Vol%	0.5 Vol%
MMT (100 µm-cell)	0 – 10000 Vol%	4 mg/l	@ 80 mg/l	20 mg/l
CMT (100 µm-cell)	0 – 10000 Vol%	7 mg/l	@ 80 mg/l	20 mg/l
Manganese (by MMT)	0 – 2500 mg/l	1.3 mg/l	@ 80 mg/l	5 mg/l

Table 5.1.b

Parameter	Range	Repeatability	at:
RON	70 – 110	0.4	@ 92 RON
MON	60 – 100	0.3	@ 82 MON
AKI	65 – 105	0.3	@ 87 AKI
IBP	20 – 50	1.8	@ 37 °C
T10	30 – 70	3.3	@ 67 °C
T50	65 – 130	5.5	@ 95 °C
T90	140 – 190	6.2	@ 150 °C
FBP	160 – 220	5.2	@ 195 °C
DVPE	45 – 100	2.8	@ 60 kPa
Density	0.5 – 1.999 gcm ⁻³	0.0006 gcm ⁻³	@ 0.7800 gcm ⁻³

Note 1: The range and repeatability of the calculated parameters (RON, MON, AKI) depend on the database used.

Note 2: The repeatabilities listed were measured with one scan per measurement. The repeatability can be improved by averaging two or more scans per second.

5.2 Diesel analysis

ERASPEC, for diesel sector works quite similarly as in gasoline and follows exactly the requirements of the European Regulation EN 14078 and measures the following parameters:

Concentrations:

Biodiesel (FAME) (ASTM, EN14078), Polynuclear Aromatics (PNA), Aromatics Cetane Improver (2-EHN).

Distinction between FAME (Biodiesel) and vegetable oil: Separate calibrations for different FAME types.

Density 0.500 – 1.999 g/cm³ (There is internal densitometer of the Anton Paar firm according to IP 559, correl. ASTM D 4052).

Properties:

Cetane number (correlation to ASTM D 613)

Cetane Index (correlation to ASTM D 976 and ASTM D 4737)

Distillation points (correlation to ASTM D 86): IBP, T10, T50, T85, T90, T95, FBP
+10 Freely programmable parameters

All the above are summarized in the following tables, where the detection limits and repeatability are shown. *Table 5.2.a* and *Table 5.2.b*. [13]

Table 5.2.a

Concentration of:	Range	Standard Deviation	Repeatability	at:	Detection Limit
Biodiesel (FAME, FAEE)	0.2 – 65 Vol%	0.07 Vol%	0.2 Vol%	0 – 10 Vol%	0.08 Vol%
		0.12 Vol%	0.3 Vol%	10 – 30 Vol%	-
Total Aromatics	0.5 – 60 Vol%	0.13 Vol%	0.3 Vol%	@ 20 Vol%	0.5 Vol%
Polynuclear Aromatics (PNA)	0.5 – 50 Vol%	0.03 Vol%	0.1 Vol%	@ 5 Vol%	0.5 Vol%
Cetane Improver (2-EHN, IPN)	50 – 20000 mg/l	9 mg/l	23 mg/l	0 – 800 mg/l	50 mg/l
Dimethoxymethane (Methylal)	0 – 20 Vol%	0.01 Vol%	0.02 Vol%	@ 0.4 Vol%	0.06 Vol%
		0.06 Vol%	0.18 Vol%	@ 2.5 Vol%	-
		0.2 Vol%	0.6 Vol%	@ 20 Vol%	-

Table 5.2.b

Parameter	Range	Standard Deviation	Repeatability	at:
Cetane Number	20 – 80	0.32	0.8	@ 50
Cetane Index	20 – 80	0.25	0.7	@ 50
T90		0.4°C	1.0°C	@ 330°C
Density	0.5 – 1.999 gcm ⁻³	0.0002 gcm ⁻³	0.0006 gcm ⁻³	0.7800 gcm ⁻³

Note 1: The range and repeatability of the calculated parameters (Cetane Number, Cetane Index, T90) depend on the database used.

Note 2: The repeatabilities listed were measured with one scan per measurement. The repeatability can be improved by averaging two or more scans per measurement.

6. Database

It is essential before analyzing anything to make clear that the section of the device for diesel works completely the same way as for gasoline. However for both sections different cells (in the interferometer), different memory sections and different database are used. In other Eraspec instruments there is the analysis option "Jet Fuel" or "Synthetic Fuel" or "Methanol Fuel". These options are selected by the customers according to their needs. The Eraspec instrument of the NTUA Laboratory for fuels and lubricants, offers the option for the analyses "Gasoline" and "Diesel". The choice is made by the user as soon as the instrument has been activated and stabilized. Details of the operation are given in the following chapters. The calibration was conducted in both sections, therefore whatever is said and referred to for one of them, fully applies for the other as well.

6.1 Libraries

The database for gasoline is structured in "LIBRARIES". The instrument has 3 libraries for each sector. Each library contains gasoline spectra the peaks of which have matched with their measured standard properties. Based on these spectra the properties of unknown samples are predicted. The software uses statistical methods described in the theoretical part to better match with the library spectra. The libraries of the instrument contain hundreds of spectra from all over the world, which have been entered by the factory. Nevertheless the manufacturers highly recommend the enrichment of the libraries with spectra of local fuels because they will be closer to the unknown samples they will face during the measurements. This makes sense because no matter how much fuels resemble to each other, the supplies of the refineries in a northern country are different from the supply of crude in EL.PE. Moreover, even in the same refineries of a country sometimes it is possible the supply to change and therefore the composition of crude. All these factors influence the composition of the final product. Thus the enrichment with new libraries from domestic refineries may improve results and make the instrument more effective. That is, the ideal library for Greece would be the one that would have been made with samples of known properties from the EL.PE and MOTOR OIL HELLAS refineries.

Nevertheless, because in order for someone to get samples with certified properties from the refineries which is a time consuming process, calibration is considered to be of great importance and the instrument provides a very user-friendly software which makes the whole procedure even more facile.

6.1.1 Library enrichment procedure

For an analyst to enter new data in the libraries of Eraspec, a sequence of events must take place. First he has to become familiar with the instrument and be trained to conduct measurements. This also includes individual settings such as repeatability and safety levels. Immediately after the user learns how to make new libraries and also which properties are demanded by the instrument for the calibration. So, in the last and most time consuming step, the user will know which properties should be stated on the certificates of the refineries in order to enter them in the libraries.

6.1.2 Starting the instrument and the reference measurement

It should be noted that regardless of which method is selected, diesel or gasoline, the measurement is conducted in exactly the same manner. The only thing that changes is the cell of the interferometer being used. Obviously there is differentiation in the databases used and in the results, i.e. with the change of method some other section of the instrument's memory is in use.

The instrument is switched on by pressing the start button. It will load the software. During this process, the LED on the front screen will flash. Loading software takes about 35 seconds. After that, the instrument remains inert until the temperature of the Laser operation is stabilized.

When the software is activated, the regulator of Laser is activated (the regulator is a software as well which depending on the circumstances regulates the movements of the Laser), and the message "wait for laser controller" appears on the screen. Because the instrument has just started, it automatically receives a reference spectrum "Reference". This is nothing more than a measurement of the air present in the cells in order for the instrument's sensors to be adapted. The controller thus orders the preheating of the Laser (to carry out the reference measurement) and the display is:

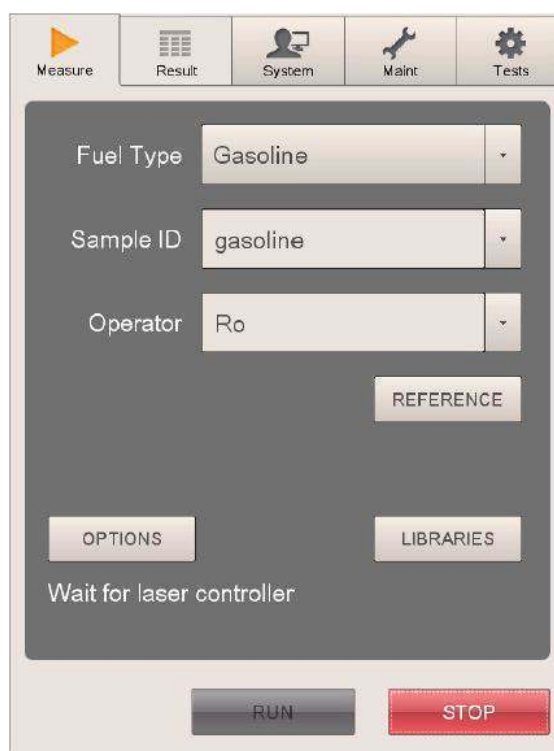


Fig.6.1.2.a

After this, the device automatically will load all available reference spectra. "Reference". If the most recent report available in the memory is more than one hour (in hours of operation) a measurement of a new reference spectrum starts automatically. More about the "Reference" are listed below. The display is:

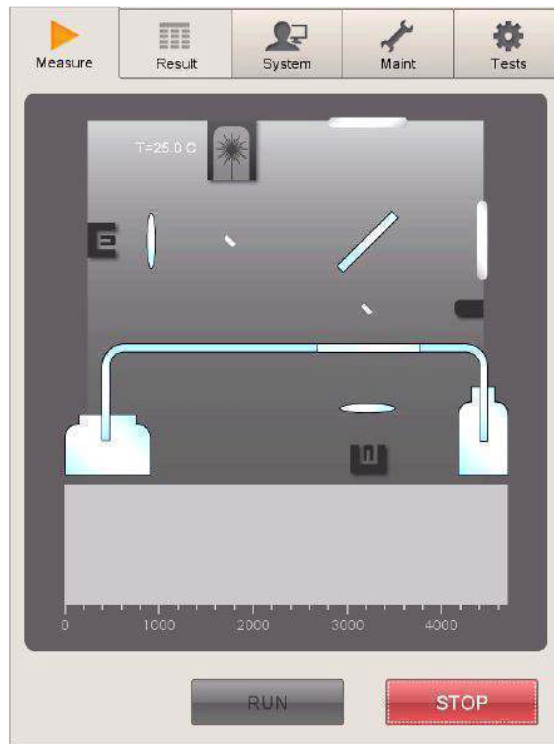


Fig.6.1.2.b

Then the source of infrared and Laser are activated:

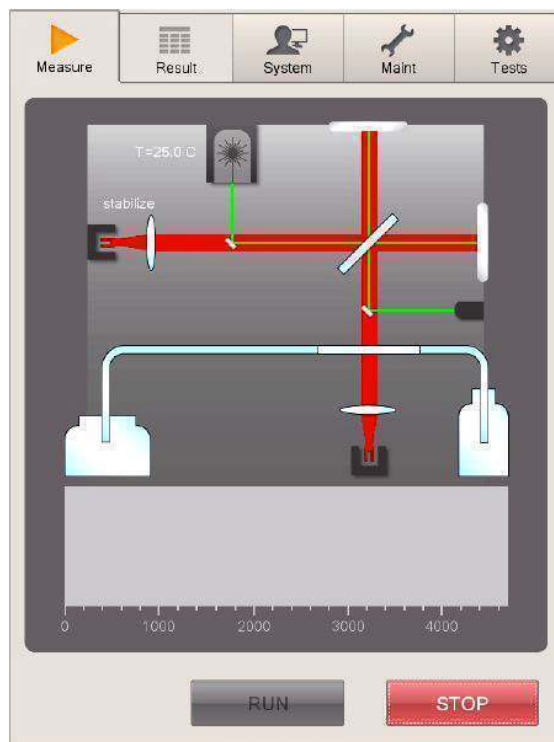


Fig.6.1.2.c

The scanning for measurement of the spectrum starts:

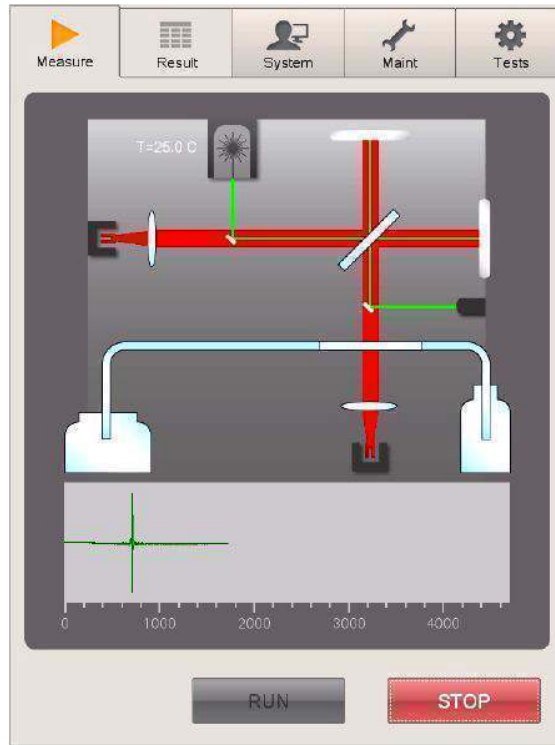


Fig.6.1.2.d

After measuring the "reference sample", the instrument is ready for measurement.

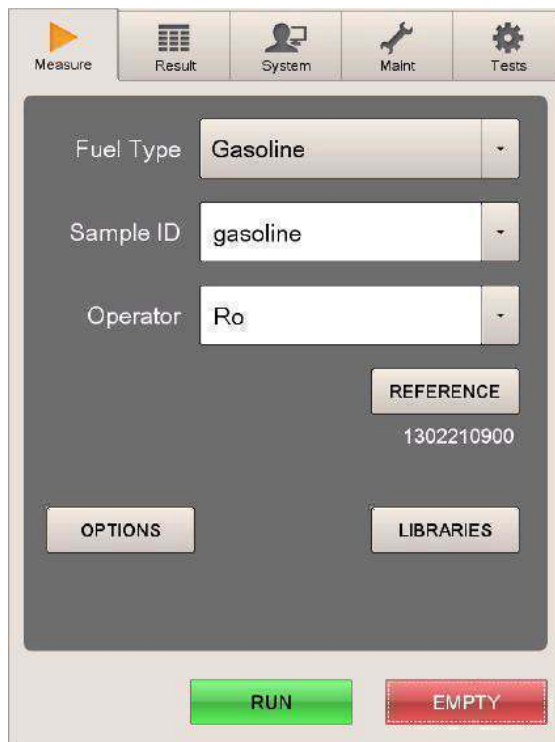


Fig.6.1.2.e

The above Figure is the original display and the menu of the instrument software.

Field "Fuel Type": selects the type of fuel for analysis. This instrument has installed software for gasoline and diesel.

Field "Sample ID": the user enters the name that later will appear in the results of the measurements he will perform.

Regarding the "Reference": Eraspec needs a reference spectrum in order to conduct the measurements accurately. The reference spectrum is essentially a spectrum obtained without sample. The spectrum obtained from a sample during the course of a measurement is then subtracted from the reference spectrum. This way any influence on the spectrum which is not due to the sample (such as the lines of the water vapors and carbon dioxide, or the influence of the dependence of the wavelength of the source of infrared emission) is removed so that only "clean" absorption spectrum of the sample is obtained which later is analyzed to determine concentrations and calculate properties.

In normal mode, the reference spectrum is automatically exported in the following cases:

- 1) After the instrument is switched on and the last reference spectrum stored in the memory is longer than one hour.
- 2) During normal operation, when the last report is more than an hour.
- 3) During normal operation, when the temperature of the spectrophotometer changes by 1° C (1.8°F) from the last measurement of a reference spectrum.

The numbers appearing under the button (i.e. 1302210900) mean "YYMMDDhhmm", i.e. year, month, day, hour and minute that the reference measurement was conducted.

The option "Libraries": It is examined below and is the libraries used by the instrument to output the results. Aim of this work is to enrich the instrument with new libraries which will contain as many as possible fuel spectra with properties similar to those of the test samples.

From this menu the user can see which libraries are in use and which are available. In addition he can see what are the spectra saved in the libraries and by whom user. Depending on the user rights of each user, deleting spectra and even entire libraries is possible.

The option "Options": The user selects how many scans should be done per measurement and also how many times the pump will be used in order for a measurement to be done. This way the accuracy of the results and the signal-to-noise ratio (SNR) described above in the theory, increase.

Eraspec usually takes spectral data during a scan and calculates the spectrum and all the results based on this scan. However, to improve the SNR of the spectrum and hence the repeatability, it is possible to make more than one scans for one measurement. From the scans the average will come out thus reducing the noise in the data. Ideally, taking the average of N spectra, the noise will be reduced by a factor \sqrt{N} (square root of N). Each scan lasts about 18 seconds, so for every additional scanning the time of the measurement increases by 18 seconds.

The maximum number of scans is 8, but by the manufacturer it is recommended maximum use of four scans for one measurement.

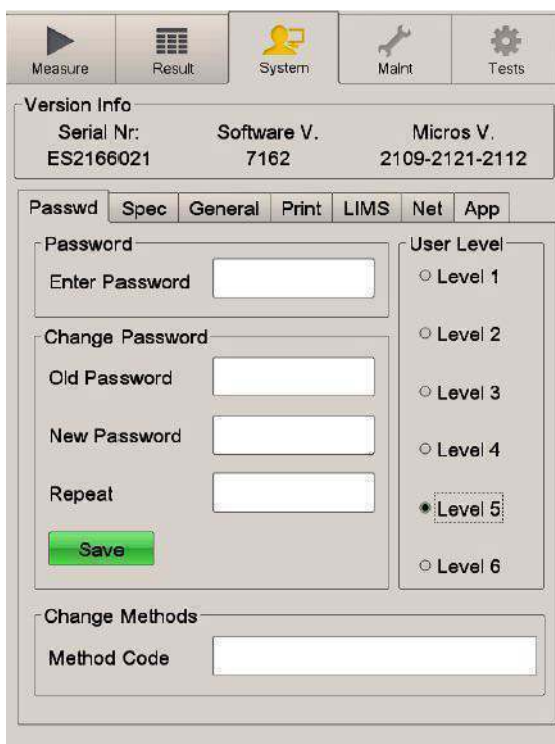
Field "Operator": Its usefulness has to do with the clear distinction between users and their rights with regards to the access of each one in the instrument. Users' levels are 6 with serial degree of accessibility. In order for a user to increase his degree he must know and enter a specific code. More specifically:

User level 1: It is the lower level with more restrictions. At this level the user can only change the sample ID and the user's name. He can make sample measurements, get reference samples, can see and print results, but also to save them and get in a USB stick. All other parameters are not accessible or even not visible and therefore can not be modified. This step allows completely unskilled users to make measurements without the risk of making any mistake.

User level 2: Allows the same functions as in level 1, and additionally, samples with known properties may be added to the database. This is an indication that highlights the importance of continuous calibration and enrichment of the instrument library with spectra of known properties.

User level 3: Allows the same functions as in level 2. In addition, all functions in the libraries can be done, and the results lists can be configured by selecting "Column" in the results menu.

User level 4: Allows the same functions as in level 3, and further, the parameters shown in the menus "Ports" and "General" can change. Also accessible at this level are the settings in the results menu.



User level 5: Allows the same functions as in Level 4, plus the code can be changed (*Fig.6.1.2.f*).

User level 6: Allows the same functions as in level 5, and further gives full access to the menu "Maint" (maintenance and service menu) which normally is not displayed, and to the menus "Network", "App" and "Test".

It is very important and must be noticed that never must parameters or settings be changed in the menu "Maintenance" without the supervision of the local representative or of Eralytics. Changing parameters may cause damage to specific elements or make the correct measurements impossible. In addition one should never use the sixth level for routine measurements with Eraspec. The menu of safety levels is shown in *Fig.6.1.2.f*.

Fig.6.1.2.f

6.1.3 Start of a normal measurement

The user being in the **main screen** (*Fig.6.1.2.a*) and all the previous steps have been done, he is now ready to make the first measurement. Having a sufficient amount of sample (over 15 ml) places the tube in the sample and presses the "RUN" key either the button or on the touch screen.

Before conducting any measurement the user must ensure that the waste container is in the correct position and not completely full.

If desired, the user can send air to pass through the system by pressing the "EMPTY" key. The same way the user also can "flush" the system if the nature of the samples he analyzes is changed completely. If e.g. he was measuring gasoline and now needs to pass to diesel, he should make 3-4 "rinses" before proceeding to measurements.

A measurement can be stopped at any time by pressing "STOP" (either the button or on the touch screen).

After the command "RUN" is pressed, the screen changes showing the sequence in which the measurement is being made.

First the sample is automatically sucked (filling process) by means of the pump inside the instrument. When this process is completed, the instrument asks the user to remove the tube from the sample. The tube must be placed in an empty bottle to avoid fumes in the laboratory.

Simultaneously with the filling process, the infrared source is activated, and then it takes a few seconds until the stabilization of the source and the infrared detector. After that, it starts the data scan as described in the beginning the spectroscopic measurement.

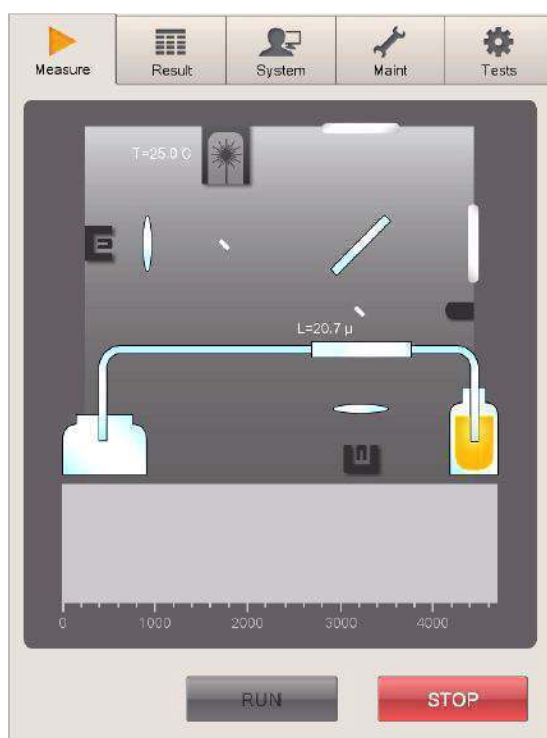


Fig.6.1.3.a

Finally the sample is removed from the system and discharged into the waste container. Together with the removal of the sample, the spectrum and the results are being calculated. These calculations last only 10 seconds. When the measurement is finished the results are displayed and saved in the results memory.

The whole measurement is fully automated and takes about 60 seconds. It shows huge similarities with the reference measurement.

The sequence of events such as shown by the instrument and as described above during the course of a standard measurement is as follows:

- Preparing the process (*Fig.6.1.3.a*).

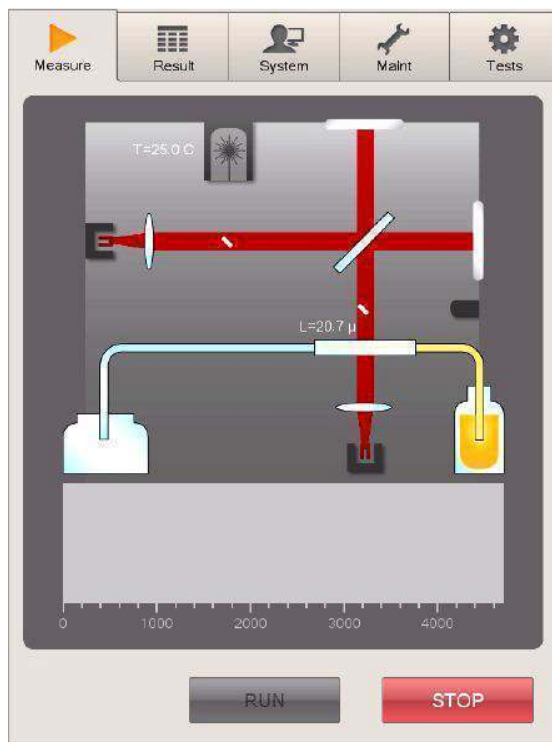


Fig.6.1.3.b

- During the filling process, the infrared source is activated (Fig.6.1.3.b).

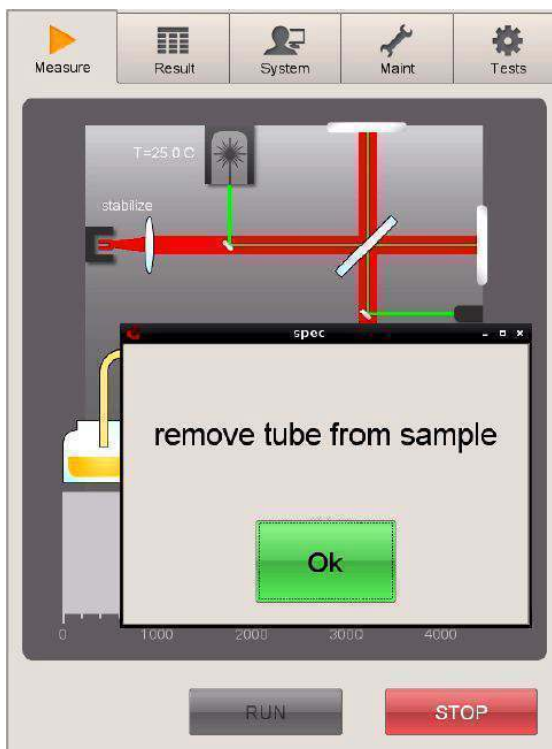


Fig.6.1.3.c

- The instrument asks the user to remove the tube from the sample (Fig.6.1.3.c).
- In this latter case the user is good to remove the tube while pressing the OK button. Moreover there should exist an empty sampling bottle nearby so that the pipe be placed there. If the pipe is not removed and a reasonable time passes by, about 30 seconds, then Eraspec continues the movement of the interferometer normally. This is not good of course since it can give distorted results

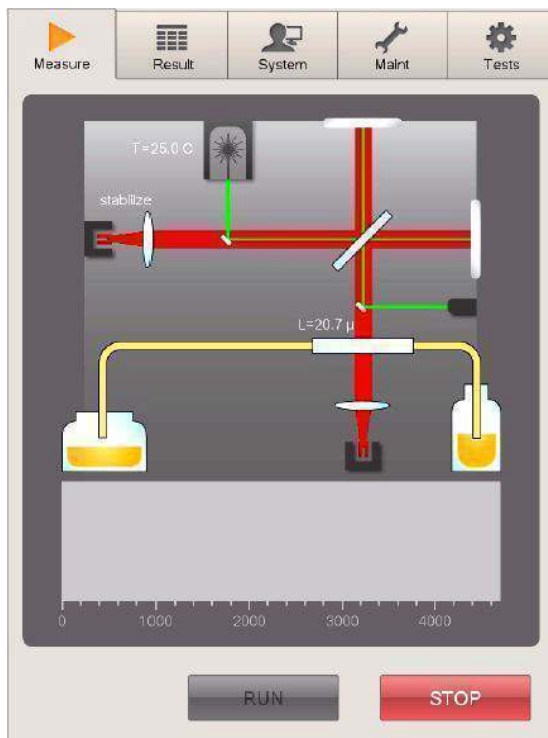


Fig.6.1.3.d

- Immediately after is activated the Laser (Fig.6.1.3.d).

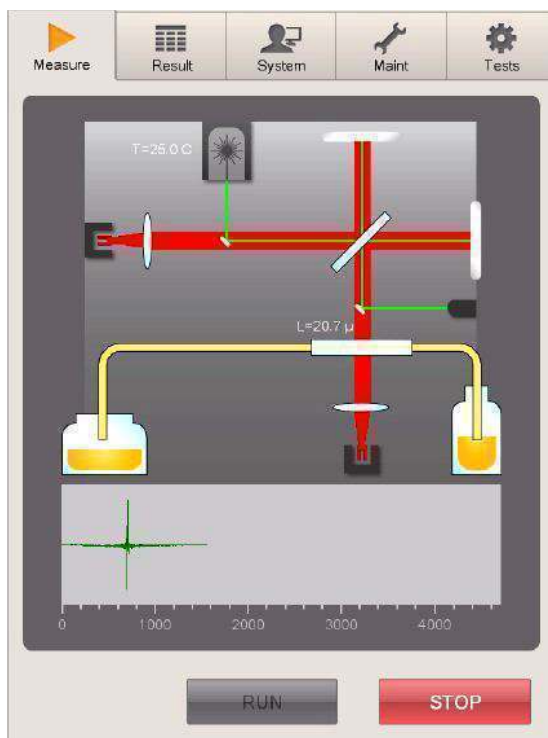


Fig.6.1.3.e

- After a few seconds, when all sources and sensors are stabilized, scanning for data acquisition starts. The sensor signal appears live at the bottom of the screen (Fig. 6.1.3.e).

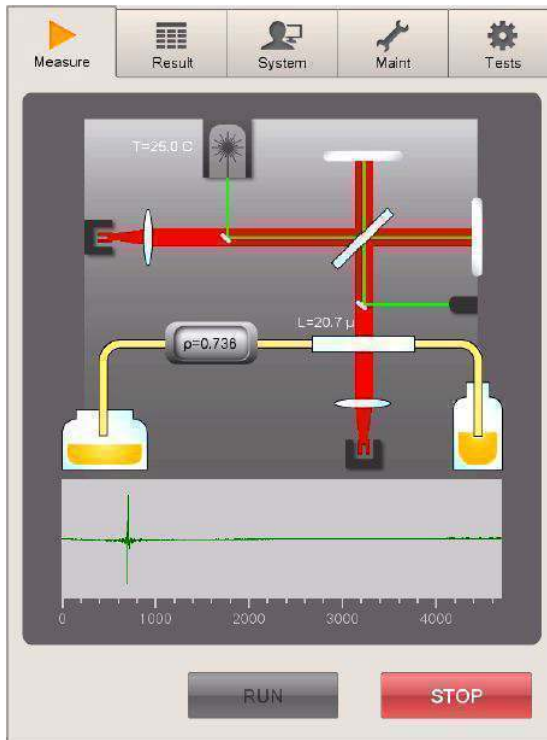


Fig.6.1.3.f

- When data collection is complete, the pump starts to remove the sample from the system (Fig. 6.1.3.f).

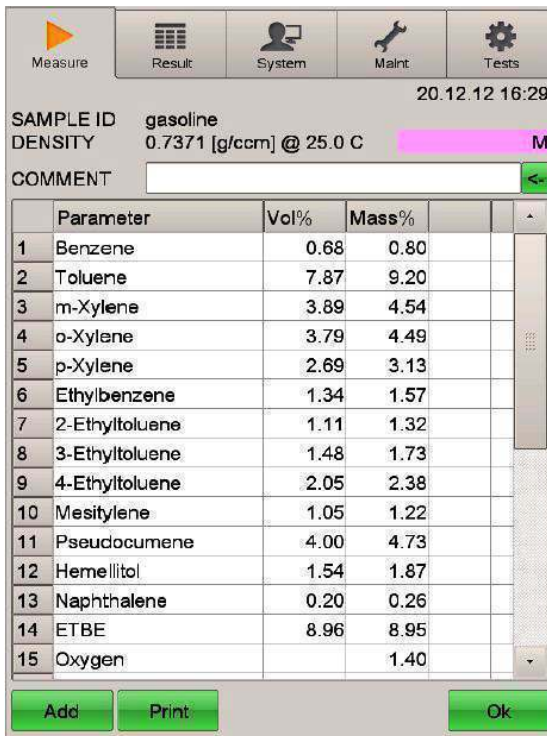


Fig.6.1.3.g

- Usually, even when the pump is still working, the list of results is displayed on the screen (Fig. 6.1.3.g).

	Time	Date	Sample-ID	RON	MON	Benz
1	17:41	20.12.12	gasoline	96.9	86.4	0.6
2	17:40	20.12.12	New solvent			
3	17:36	20.12.12	New solvent			
4	17:15	20.12.12	gasoline	95.9	85.4	0.6
5	16:29	20.12.12	gasoline	97.0	86.1	0.6
6	15:44	20.12.12	gasoline	94.9	86.7	0.2
7	15:43	20.12.12	gasoline	95.3	87.0	0.1
8	14:45	20.12.12	Test	95.6	87.2	0.1
9	14:42	20.12.12	Test	95.7	87.2	0.1
10	14:10	20.12.12	Test	95.8	87.2	0.2
11	14:03	20.12.12	Test	95.7	87.1	0.2
12	13:54	20.12.12	gasoline	95.4	87.0	0.2
13	13:46	20.12.12	gasoline	95.5	86.9	0.1
14	12:24	20.12.12	gasoline	94.9	86.5	0.1
15	12:15	20.12.12	gasoline	123.9	105.8	

Navigation buttons: Measure, Result, System, Maint, Tests. Control buttons: Save, Add to Lib, Print, Edit, Result, Column.

Fig.6.1.3.h

- Otherwise at the end of the measurement again the main screen is displayed and the user manually goes to the results menu. That is, he presses the "Result" key and the image at the left is displayed (Fig. 6.1.3.h).

Either one or the other case, it would be the right time for the sample under measurement to be recorded in the libraries. Thus, we call this point, **point A** and later we will return to that. The explanation of creating a new library is preceded, and the next chapter refers just to that.

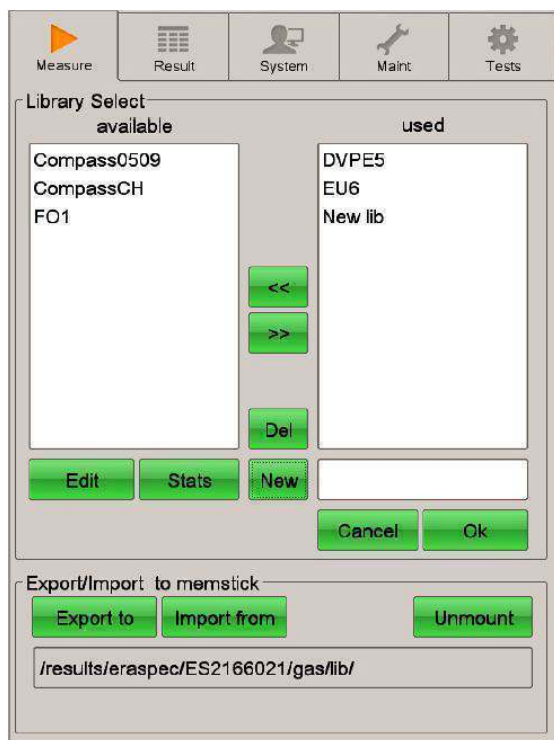


Fig.6.2 The libraries menu

6.2 Creation of a new library

From the main menu: "Libraries" > "New" the user enters the name of the future library and then confirms it by pressing "New" (Fig.6.2). At previous time he has created a new library, within which he will enter the new samples. He also might enter the samples into an existing library, that is to enrich it. However it is more correct to separate libraries depending on the samples they contain i.e. to be grouped. Other library for EL.PE, other for MOH, and others for the factory. This must occur because if all samples were entered in one library, the user would be not able to remove some of them, depending on the analysis he wants to conduct. It would be unavoidable to take into account all the spectra. Also in this instrument there has been created a new library with the name "New Samples" in which members of the laboratory will enter samples of known

properties if they come in their possession.

6.2.1 Input of samples in the libraries

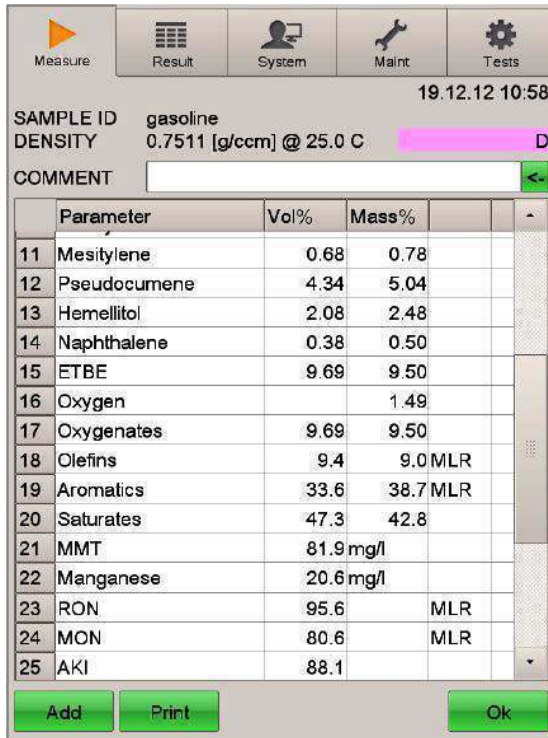


Fig.6.2.1.a

Let us return now to the time **point A** of section 6.1.3 where immediately after the measurement the results are displayed on the screen. The user simply presses the key "Add" for direct input of them into one of the libraries. Alternatively the user goes to the option: "add to lib" provided he is already in the subsection "Result" and has marked the sample for which he knows its properties.

- Direct "Add" (Fig.6.2.1.a).

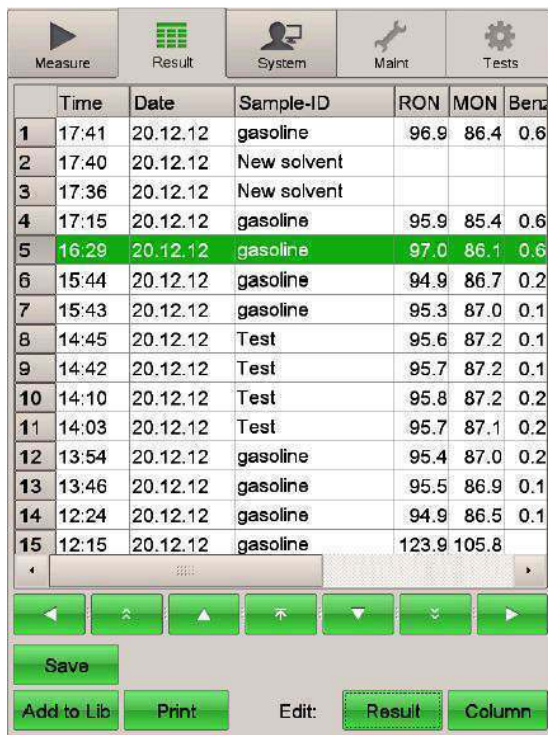


Fig.6.2.1.b

- "Add to Lib" (Fig.6.2.1.b).

Being at menu "Result" and after having marked the desired sample, press the "Add to lib" or if "Add" had been used, a list of all available properties calculated with the use of the database appears on the screen.

Parameter	Value	Unit
1 RON		
2 MON		
3 Aromatics		Vol%
4 Olefins		Vol%
5 IBP		C
6 T10		C
7 T50		C
8 T90		C
9 FBP		C
10 DVPE		kPa
11 E300		V%
12 DVPE		kPa
13 FREE04		
14 FREE05		

Fig.6.2.1.c Gasoline

Parameter	Value	Unit
1 Cetane Number		
2 Cetane Index		
3 Cetane Improver		mg/kg
4 Aromatics		Vol%
5 PNA		Vol%
6 IBP		C
7 T10		C
8 T50		C
9 T85		C
10 T90		C
11 T95		C
12 FBP		C
13 E250		V%
14 E350		V%

Fig.6.2.1.d Diesel

The left figure concerns Gasoline and the right one Diesel. In these windows the user enters the values from the refineries and confirms by pressing "OK". Values with decimal points with comma (,) are entered in the system with a corresponding dot (.) In case of entering comma, the instrument halts and switches off automatically, resulting in major loss of time.

It is obvious that great emphasis is given to distillation temperatures, because most conclusions for other properties of the measured samples as well, depend on them.

This way, the peaks of the spectrum obtained by the measurement are assigned to the certified properties. Therefore this spectrum is "baptized" on the basis of the standards and is now available in the library. For maximum accuracy, it is advised that many samples from the same bottle to be entered in the library more than once. This reduces even more the potential error of the instrument. Finally this spectrum is saved in any library that the user wants. The need for more than one inputs, is understandable by errors in some values in the exported spectra of the following chapters.

6.3 Refinery protocols

As part of this work the calibration of ERASPEC with samples of known-certified properties from these two companies (EL.PE and MOTOR OIL HELLAS) was conducted.

Specifically, 30 sample bottles with gasoline and 13 bottles with diesel were received, each bottle containing approximately 100 ml of gasoline or diesel. Each bottle had a tag with a protocol number and a tank code. The document accompanying the samples contained the date and time of sampling and the tank the sample was taken from. The protocols received had distinct differences between them, i.e., some did not contain all the properties that Eraspec accepts for the calibration, while others had so little information that it was necessary to send the codes back in order for new certificates to be issued. More clearly the differences are shown and are analyzed below.

Indicative protocols were of the following formats:

790001654120	UNL2000	GASOLINE UNLEADED 95 RON			
17.05.2014	06:49:30	2401A	FINISHED	TANK	
0002 A/A ABITMATOE		0	15,000	4,728	A
0010 DENSITY AT 15 C		0,7200	0,7750	0,7292	A
0020 Research Octane Number		95,0	100,0	97,0	A
0030 MON		85,0	100,0	86,0	A
0040 Induction Period		360	1,000	420	A
0060 Lead in Gasoline		0,000	0,005	0,001	A
0062 SULFUR, ppm		0,0	10,0	7,4	A
0070 Existent Gum		0	5	2	A
0080 DVPE		45,0	60,0	60,0	A
0090 Benzene		0,00	1,00	0,87	A
0100 Evaporated at 70 C		20,0	48,0	32,0	A
0110 Evaporated at 100 C		46,0	71,0	61,0	A
0112 DISTILL, EVAP 150oC		75	0	93	A
0130 Distillation, FBP		100,0	210,0	186,0	A
0140 Distillation Residue		0,0	2,0	1,2	A
0160 AROMATICS		0,0	35,0	16,4	A
0180 Oxygenates Content		0,00	15,00	5,60	A
0200 Copper Corrosion				1A	A
0210 Oxygen		0,0	2,7	1,1	A
0220 Olefins		0,0	18,0	18,0	A
0230 DENS AT 15 C UPPER		0,0000	1,0000	0,7292	A
0240 DENS AT 15 C MED		0,0000	1,0000	0,7292	A
0250 DENS AT 15 C BOT		0,0000	1,0000	0,7292	A
0260 DVPE UPPER		45,0	60,0	60,0	A
0270 DVPE MED		45,0	60,0	60,0	A
0280 DVPE BOT		45,0	60,0	60,0	A
790001654191	UNL2000	GASOLINE UNLEADED 95 RON			
17.05.2014	15:06:27	2401A	TANK CORRECTION		
0002 A/A ABITMATOE		0	15,000	4,843	A
0010 DENSITY AT 15 C		0,7200	0,7750	0,7296	A

Fig.6.3.a Gasoline Sheet, Cetane number 95 R, from MOH (Motor Oil Hellas)

In the protocol depicted above (Fig.6.3.a) are listed:

- 1) Top left corner, the serial number
- 2) Immediately below, the date of sampling
- 3) "UNL2000" is the code of the tank and immediately below the exact time of sampling
- 4) "GASOLINE UNLEADED 95 RON" is the category of gasoline based on the octane number
- 5) Code 2401A is the sample code, marked on the bottle

In total 15 such samples were received, 3 of which were with wrong certificates and therefore not used. From the 12 correct samples, their codes and the corresponding sampling dates were:

<u>Date</u>	<u>Code</u>	<u>Octanes</u>
14/4/2014	2401A	100
15/4/2014	2401B	95
20/4/2014	2401A	95
28/4/2014	782	95
1/5/2014	785	95
13/5/2014	2401A	95
13/5/2014	8107	95
14/5/2014	2401B	95
16/5/2014	2401B	95
17/5/2014	2401A	95
20/5/2014	2401B	95
22/5/2014	704	95

Properties	Units	Results	Methods
⁽¹⁾ Density at 15degC	kg/m ³	831.8	EN ISO 12185
Colour		L0.5	ASTM D1500
Appearance		Clear and bright	ASTM D4176
% (w/v) Recovered at 250degC	% w/v	26.5	EN ISO 3405
% (w/v) Recovered at 350degC	% w/v	92.9	EN ISO 3405
95% (w/v) Recovered at degC	degC	355	EN ISO 3405
⁽¹⁾ Flash point	degC	66.0	EN ISO 2719
⁽¹⁾ Sulphur content	mg/kg	3.7	EN ISO 20846
Copper strip corrosion (3hrs at 50degC)		class 1A	EN ISO 2160
⁽¹⁾ Cold Filter Plugging Point (CFPP)	degC	-10	IP 309
Cloud Point	degC	-6	EN 23015
Fatty Acid Methyl Ester (FAME)	% w/v	Nil	EN ISO 14078
Viscosity at 40degC	mm ² /s	3.189	EN ISO 3104
Water content	mg/kg	71	EN ISO 12937
Cetane number		54.6	EN ISO 5165
Cetane index		57.2	EN ISO 4264
Ash content	% m/m	0.003	EN ISO 6245
Carbon residue (on 10% distillation residue)	% m/m	0.01	EN ISO 10370
Total contamination	mg/kg	5.5	EN 12662
Oxidation stability	g/m ³	3.5	EN ISO 12205
Oxidation stability	hours	21.0	EN 15751
Poly cyclic aromatic hydrocarbons	% m/m	2.4	EN 12916
Lubricity, corrected wear scar diameter	µm	443	EN ISO 12156-1
Conductivity at 20degC	pS/M	220	ISO 6297

Fig 6.3.b Certificate for Diesel sample, from MOH

Regarding the above certificate (Fig.6.3.b) the paper was torn so that the rest of the data are not shown. When it was received was accompanying the sample bottle. Similarly, the next certificate (Fig.6.3.c) was truncated and accompanying the sample bottle.

Density at 15 oC Kg/l	0,8323
Colour ASTM	L0,5
Appearance	clear bright
Flash point °C	66,0
Kin Viscosity cst at 40°C	3,229
Pour point °C	
Cloud point °C	-6,8
C.F.P.P. °C	-9
Sulphur ppm	5,0
Sulphur wt%	
Copper corrosion	1a
Distillation	
IBP oC	172
10% oC	221
20% oC	243
50% oC	284
65% oC	303
85% oC	335
90% oC	345
95% oC	358
FBP oC	363
Vol % at 250 °C	23,9
Vol % at 350 °C	92,3
Vol % at 360 °C	95,7
Cetane index	57,7
Cetane number	
Hydrogen Sulfide H ₂ S mg/kg	
Micro carbon residue wt%	
Ash wt%	
Total acid number mgKOH/g	
Oxid. stability g/m ³	
Oxid. stability accel. hours	
Water content ppm	82
Water and sediment %vol	
Lubricity (wsd 1.4) µm	
Separability min	
Quinizarine ppm	
Quinizarine test	
Solvent yellow 124 ppm	
Solvent yellow 124 test	Neg
Elect.conductivity Ps/m	43
Polycyclic aromatic H/C wt%	2,7
FAME content % v/v	Nil

Fig.6.3.c Certificate for Diesel sample, from MOH

Only the characteristics of Olefins and Aromatics relating to instrument calibration were recorded on the next sheet as well as on the rest that were taken from the same batch of samples. Therefore these codes were gathered and sent back to the refineries in order for them to send correct certificates with the rest of the necessary properties of the same samples. Totally, the samples of this type were 10, and their codes are shown below.

The code here is the number of handwritten characters, i.e. 2286192.

reformulyzer		763a		AG			
Data FileName	C:\Chem32\1\DATA\ID1405\763a.D						
Operator	Christina		Vial	: 101			
Acquired On	12-Jun-14 10:52:03 AM		Inj Volume	: 0,1 µl			
Processed On	: 12-Jun-14 10:52:06 AM						
Sample Name	: 763a						
Sample Group	: Winerspec MTBE/TAME						
Sample Type	: Winerspec MTBE						
Instrument Method	: Winero1						
2286192							
Analysis Conditions							
A	3	EthAlcSep	110				
B	3.85	OlefinSep1	100				
E	2.5	OlefinSep2	150				
Sample Density		742.3 kg/m ³ (at 20 °C)					
Normalized weight percent results							
C-nr	Naph.	Paraf.	Cycl Ol.	Olef.	Arom.	Oxyg.	Total
1							
2							
3		0.11					0.11
4		1.80		0.83			2.63
5	0.48	13.36	0.18	2.80		3.30	20.10
6	2.53	11.64	0.60	2.70	0.93	4.36	22.76
7	1.48	5.23	0.73	1.23	10.49		19.17
8	0.98	1.80	0.36	0.62	14.77		18.52
9	0.45	0.81	0.11	0.26	11.21		12.84
10	0.12	0.38		0.10	2.45		3.04
11+		0.21			0.62		0.82
Poly							
Total	6.82	35.33	1.99	8.55	40.47	7.65	100.00
Normalized volume percent results							
C-nr	Naph.	Paraf.	Cycl Ol.	Olef.	Arom.	Oxyg.	Total
1							
2							
3		0.16					0.16
4		2.31		1.02			3.33
5	0.46	16.20	0.17	3.21		3.30	23.34
6	2.46	13.14	0.57	2.95	0.78	4.20	24.11
7	1.44	5.68	0.69	1.31	8.99		18.08
8	0.94	1.88	0.34	0.63	12.63		16.42
9	0.42	0.82	0.10	0.27	9.56		11.18
10	0.11	0.37		0.10	2.08		2.67
11+		0.20			0.52		0.72
Poly							
Total	6.82	40.75	1.88	9.49	34.68	7.60	100.00
Oxygenates results							
Oxygenate	WT%	LV%	O2%				
MTBE	3.25	3.26	0.50				
TAME	4.36	4.20	0.68				
tert-amyliccohol	0.05	0.04	0.01				
Total	7.65	7.50	1.28				
EN ISO 22854 / ASTM D 6839							
Benzene	0.78 LV%						
Total Aromatics	34.68 LV%						
Total Olefins	11.37 LV%						
Total Oxygenates	7.50 LV%						
Total Saturates	46.97 LV%						
Total Oxygen	1.28 WT%						

Fig.6.3.d Certificate from EL.PE

On this specific sheet (Fig.6.3.d) the exact time that the sample was acquired, and also the time that it was processed are clearly recorded. This is observed on all the sheets of this type and the time difference of sampling is minimal, making the results even more reliable. The identity characteristics of these 10 samples are:

<u>Sample name</u>	<u>Code (LIMSID)</u>
763a	286192
764a 31/05/2014	2281825
763A 03_06_14	2282751
763B 10_06_14	2284179
763a	2286879
764a	2287550
711a	2288477
764A 24_06_14	2290243
763 B 27_06_14	2291620
763 A 01_07_14	2292857

The following sheet (*Fig.6.3.e*) was received from EL.PE company and in one page the certificates for five samples of gasoline are shown. In this case the name and the protocol number of the sample are displayed.

		763A	763A	764A	763B	711A
		2195363	2199313	2202199	2200869	2267691
Distillation						
IBP	°C	33.5	33.0	29.1	31.3	35.1
5	°C		40.7	37.7	40.6	42.2
10	°C	51.0	47.0	43.8	47.0	54.9
20	°C	59.2	56.4	53.1	55.6	59.9
30	°C	68.3	65.7	62.1	64.8	69.8
40	°C	78.3	73.9	74.5	74.6	81.2
50	°C	90.0	84.5	82.9	86.2	94.6
60	°C	104.7	98.7	97.7	100.2	107.7
70	°C	121.7	115.5	115.4	115.9	119.5
80	°C	137	132.8	132.5	132.2	133.8
85	°C	144.9	141.1	140.7	140.6	142.0
90	°C	153.8	149.1	147.2	148.7	151.0
95	°C	166.4	160.9	161.7	160.2	162.9
FBP	°C	190.2	185.8	178.6	183.3	181.7
Evap @ 70 °C	% v/v	32.7	35.4	36.1	35.5	30.0
Evap @ 100 °C	% v/v	57.5	61.1	61.4	59.8	54.1
Evap @ 150 °C	% v/v	89.0	90.7	90.6	91.0	89.4
Density	gr/ml	0.7434	0.7397	0.7400	0.7411	0.7528
Aromatics	% v/v	31.8	31.7	31.3	31.3	34.0
Olefins	% v/v	11.1	11.5	10.3	10.3	9.4
Ethers	% v/v	7.25	8.43	7.70	8.10	14.68
MTBE	% v/v	3.79	4.76	4.28	4.41	8.9
TAME	% v/v	3.46	3.67	3.42	3.69	5.77
Benzene	% v/v	0.84	0.81	0.80	0.82	0.77
RVP	psi	10.20	11.28	11.30	10.60	9.90
RON		95.6	95.7	95.5	95.3	100.0
MON		85.0	85.0	85.1	85.0	88.1

Fig.6.3.e Certificate of five gasolines, from EL.PE

The identity characteristics of these 5 samples are:

<u>Sample name</u>	<u>Protocol Number</u>
763A	2195363
763A	2199313
764A	2202199
763B	2200869
711A	2267691

The sheet of the next page (*Fig. 6.3.f*) is only for a single diesel sample.

This was the most complete and gives information on the following:

- 1) Plant Code (tank): P-8714A
- 2) Date: 06/05/2014 at 11:12
- 3) Protocol number (ID): 2283821

Plant Unit : P-8714A ID : 8283824
 : ΔΕΣΑΜΕΝΗ P-8714 A

Start Date : 5-JUN-2014 11:12 Report Generated On 06/06/14
 11:12
 End Date : 06/06/14 11:12
 Shift :

	Method	Units	P8714AREP_05 0614B
Product			DIESAUTSUM
Sampled Time			08:59
MONO AROMATIC HYDROCARBO	IP 391	% m/m	22.1
DI AROMATIC HYDROCARBON	IP 391	% m/m	13.8
TRI AROMATIC HYDROCARBON	IP 391	% m/m	0.3
POLYCYCLIC AROMATIC HYDR	IP 391	% m/m	4.1
ASH	ASTM D-482	% m/m	0.005
CETANE INDEX	ASTM D-4737		53.1
CETANE NUMBER	DIN 51773		53.3
CEFP	IP 309	oC	-2
COLOR			ΦΥΣΙΚΟ
COPPER CORROSION CLASSIF	ASTM D-130		1
DEL710			1
DENSITY AT 15oC	ASTM D-4052	g/cm ³	0.8380
IBP	ASTM D-86	oC	167.4
5% RECOVERED	ASTM D-86	oC	196.1
10% RECOVERED	ASTM D-86	oC	208.7
20% RECOVERED	ASTM D-86	oC	228.1
30% RECOVERED	ASTM D-86	oC	245.9
40% RECOVERED	ASTM D-86	oC	263.9
50% RECOVERED	ASTM D-86	oC	278.7
60% RECOVERED	ASTM D-86	oC	293.3
70% RECOVERED	ASTM D-86	oC	308.1
80% RECOVERED	ASTM D-86	oC	323.5
90% RECOVERED	ASTM D-86	oC	332.0
95% RECOVERED	ASTM D-86	oC	341.5
99% RECOVERED	ASTM D-86	oC	356.4
FBP	ASTM D-86	oC	367.7
RECOVERY	ASTM D-86	% v/v	98.7
RESIDUE	ASTM D-86	% v/v	0.0
LOSS	ASTM D-86	ml	1.3
RECOVERED AT 250 oC	ASTM D-86	% v/v	32.0
RECOVERED AT 350 oC	ASTM D-86	% v/v	93.2
RECOVERED AT 360 oC	ASTM D-86	% v/v	95.8
RECOVERED AT 370 oC	ASTM D-86	% v/v	
UNIT	ASTM D-86		
MEASUR_PROGRAM	ASTM D-86		
FATTY ACID METHYL ESTER	EN 14078	% w/v	7.0
FATTY ACID METHYL ESTER	EN 14078	% v/v	17.20
FLASH POINT	ASTM D-93	oC	62.5
CORR. WEAR SCAR DIAMETER	ISO 12156	um	181
SOLVENT YELLOW 124	DIN 51626	mg/L	0.0
MICROCARBON RESIDUE OM 1	ASTM D-4530	% m/m	0.02
TOTAL INSOLUBLES	ASTM D-2274	mg/L	<25
OXIDATION STABILITY 110o	EN 15751	h	>30.0
SULFUR	ASTM D-5453	mg/kg	17.2
TOTAL CONTAMINATION	ASTM D-6217	mg/l	<24.0
KINEMATIC VISCOSITY AT 4	ASTM D-445	cSt	2.8
WATER	ASTM D-6304	mg/Kg	175.00

Fig.6.3f

7. The statistical methods in the instrument

As previously discussed, Eraspec needs a library or database containing spectra and known values for calculated properties such as octane number, distillation properties and other parameters. When an unknown sample is analyzed, Eraspec uses two different methods called "Cluster Analysis" (CA) and "Multilinear Regression" (MLR).

7.1 Results for the calculated properties: "CA"

Cluster or complex analysis is the task of grouping a set of objects in such a way that objects in the same group (cluster) is more similar (on a property) to one another from others belonging to other clusters. As just discussed in the theoretical part, depending on the grade of the differences between the patterns and the unknown spectrum they are characterized as:

Cluster analysis: CA1, CA2, CA3 and CA4

Cluster analysis is based on the following principle: The spectrum of the unknown sample as well as of any sample in the libraries being used is compared in specific spectral regions (peaks) and automatically, a very small limit is set for the difference. If the difference is less than the limit, the sample from the library is very close to that of the unknown spectrum, hence the properties of the unknown and of the corresponding in the library will be almost similar (very close). Finally, the properties of the sample in the library are given as results for the properties of the unknown sample.

If more than one sample in a library have differences greater than the limit, then the result of the properties of the unknown sample is the average value of the properties of all the samples that are above the limit.

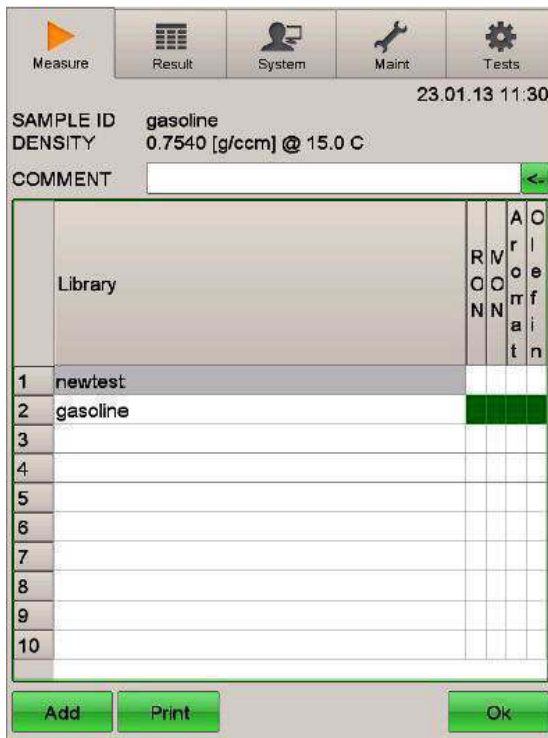
If at least one sample from the database is found under the original very low limit, then "CA1" appears in the corresponding line in the list of the results, indicating this way that at least one such sample was found during the first repetition.

If no sample based on the original very low limit is found, then the limit is doubled, and the calculation is repeated. If at least one sample from the database is found within the new limit, then the results will be displayed as described above, and "CA2" appears.

A third and a fourth repetition are performed with increasing limits and "CA3" or "CA4" are displayed after the corresponding repetition if one or more samples of the library found within the limits.

If even then there is not any sample found in the database, the cluster analysis process stops, and the MLR method is applied. The MLR is described below in detail.

When the cluster analysis is used, Eraspec can also display the samples that were found in the library, in the list of results.



Clicking on any gray box with the number at the left side in a line of some property calculated with CA, the Fig.7.1.a appears.

Fig.7.1.a

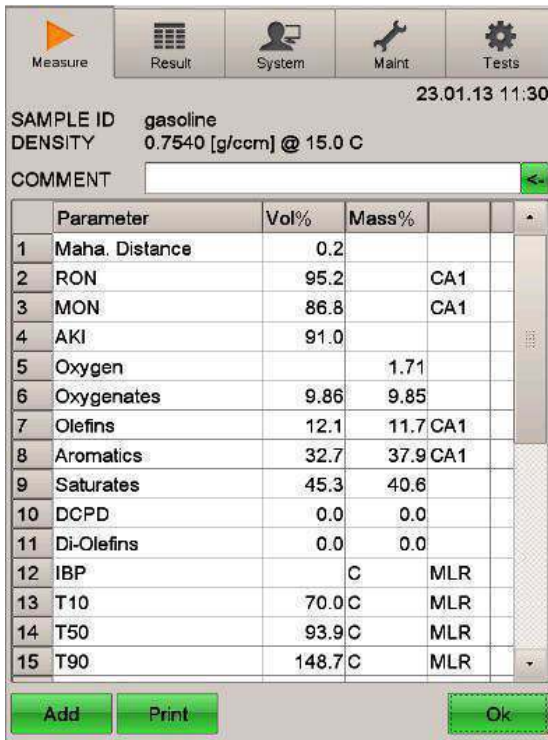


Fig.7.1.b shows a table that identifies the samples found, what libraries they were found in, and what parameters the corresponding samples were used for.

Fig.7.1.b

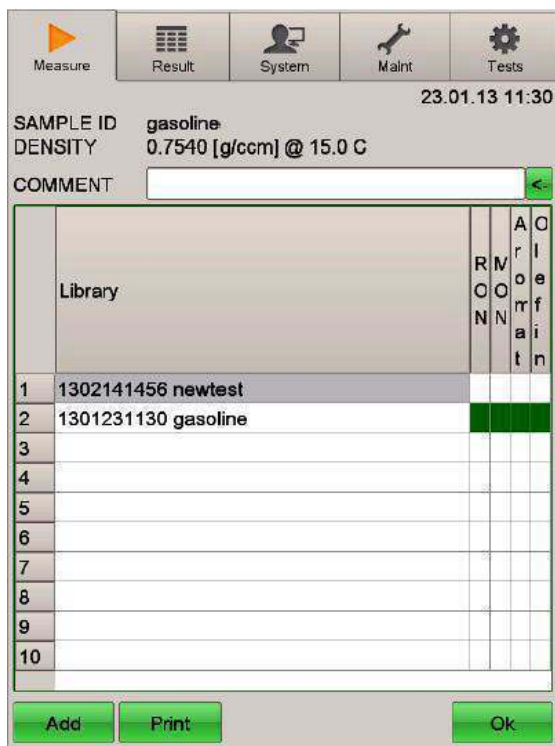


Fig.7.1.c

Sometimes the names of the samples are insufficient to distinguish which one was used, e.g. in cases that two samples have the same name. In these cases it is possible to display on the screen the date and time of measurement: click the large gray field "Library" in the header (Fig.7.1.a).

The date and time are displayed along with the name of the sample, so that each sample library is definitely recognized.

Pressing "OK" you return back to the result list.

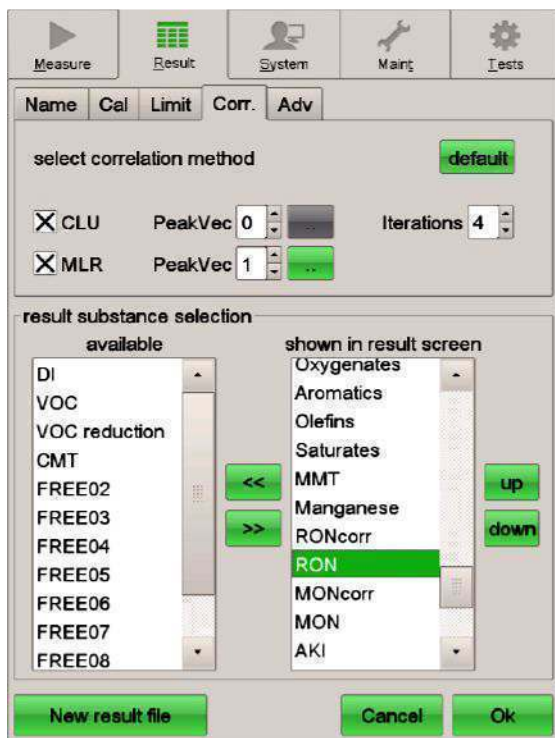


Fig.7.1.1

The number of iterations for each calculated parameter with CA, must be set separately

7.1.1 Change the number of repetitions for Cluster Analysis

In CLU line (Fig.7.1.1), the number shown in the white box next to the "iterations", indicates the number of repetitions that will occur, with increasing acceptance limits in order for a matching spectrum from the library to be found as described in section "Cluster analysis: CA1, CA2, CA3 and CA4".

Usually 4 such repetitions are made. If you need to do less or more repetitions, the user sets the desired number in the "iterations" box.

7.1.2 Choosing the peaks to use for CA

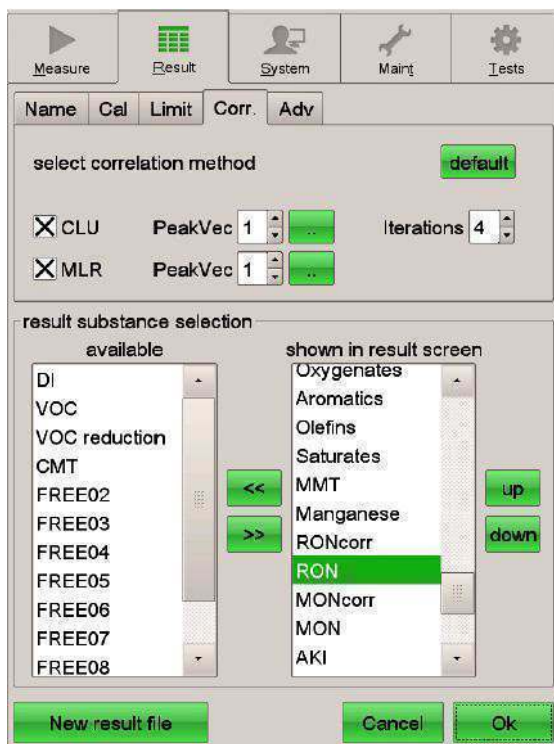


Fig.7.1.2.a

In order to find a possible match, all the peaks present in the library are used to be compared with the unknown spectrum. Only if all the peaks are sufficiently close, then the spectrum is acceptable for CA.

If for some reason, not all the peaks are used for CA it is possible to exclude one or many of them. In CLU line (Fig.7.1.2.a), the number shown in the white box next to the "PeakVec" is usually 0, which means that all the peaks are used for comparison.

To exclude peaks, the user must enter a different value in this box.

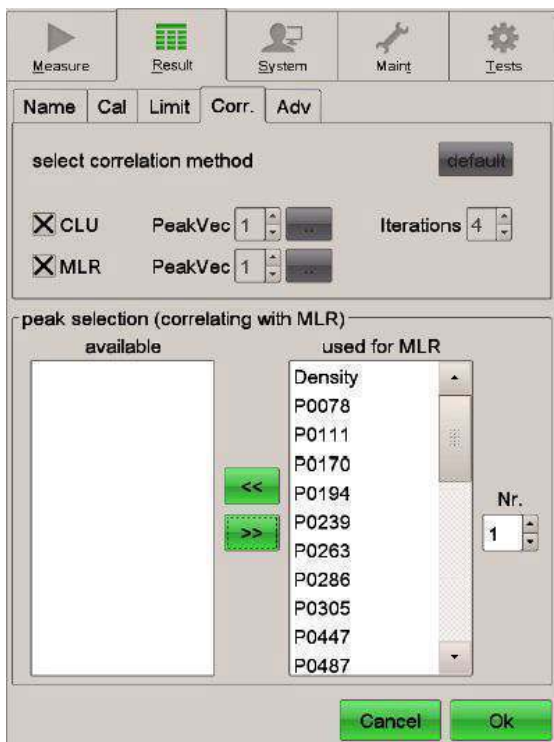


Fig.7.1.2.b Peaks for CA

Then the user touches the green box "." button next to "PeakVec" (Fig.7.1.2.a) and the list of all the peaks used for CA appears (Fig. 7.1.2.b).

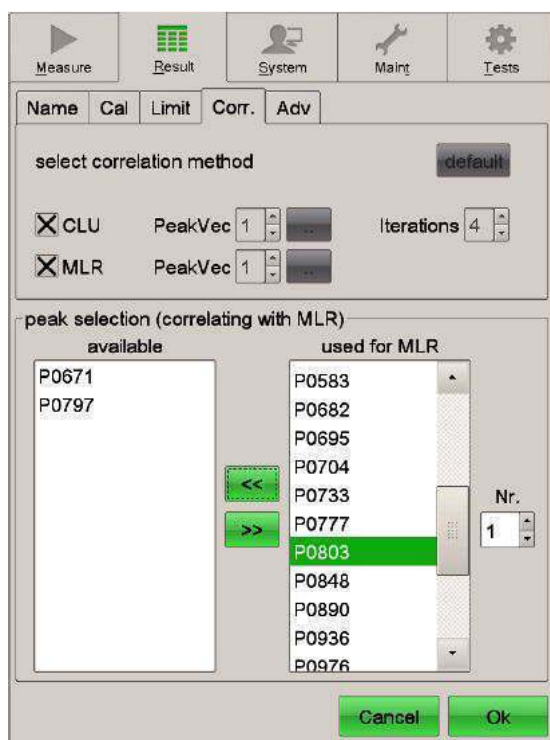


Fig.7.1.2.c

The user can move the peaks not used from the right box to the left indicating "Available" (Fig.7.1.2.c).

Press the OK key to enter this new option. The peaks that were shifted in the left box will not be used to determine whether a spectrum that exists in the library is quite matching with the unknown spectrum.

"PeacVec" for the CA must be set differently for each parameter calculated with CA.

7.2 Results for the calculated properties: "MLR"

Generally the MLR (Multiple Linear Regression) as discussed in the theoretical part is an approach for modeling the relationship between a scalar dependent variable y and one or more independent variables x (perhaps not statistically independent). In the case of one independent variable the process is called Simple Linear Regression. For more than one independent variables the method is called Multiple Linear Regression.

The MLR is a standard procedure used in many different applications in spectroscopy and other fields for data analysis.

The MLR uses the entire database (libraries) chosen for the measurement and calculates coefficients for each of the regions of the spectra used for the calculation of properties. Then from the values of the corresponding spectral regions of the unknown sample and these coefficients the properties are calculated. If the results are obtained this way, then the indication "MLR" appears for the corresponding values in the list of results.

To activate the MLR and set it to function for a property, a minimum number of 40 samples is required which must have this property being of known value. If the library or libraries used contain fewer than 40 values for a property, the MLR can not function for this property.

7.2.1 Accuracy and deviations

The CA works only if at least one sample from the database is very close to the unknown sample. The results obtained this way are usually very accurate. Of course, for the CA to work, a database with samples representative of the unknown samples which are analyzed is required. For example, most suitable of all would be a library of samples from the MOH (for unknown sample from AVIN station)

The MLR works anytime. The accuracy of the results obtained this way is usually lower than that of the CA, but if the unknown sample is not so different from samples that exist in the database, the results usually remain good.

A statistical measure of how well an unknown spectrum is represented by the database is the "Mahalanobis" distance, or in short "Maha distance". This distance is calculated and also appears in the results menu as Maha Distance. The smaller this distance, the better the spectrum is represented in the database.

What can also happen is that the unknown sample is not sufficiently represented in the database, for example, all the samples are very different from the unknown. In this case the sample is called "outlier", because it "lies outside" of the samples in the database.

Outlier (in statistics): A point (observation, experiment, measurement and so on) which lies substantially away from the other points, that is outside the range of sample libraries.

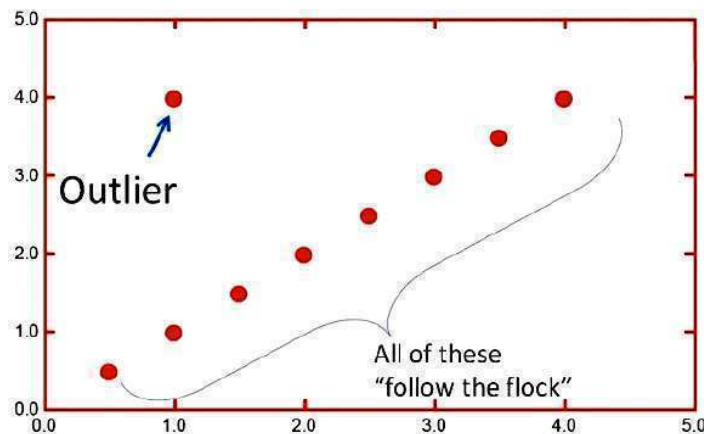


Fig.7.2.1 A point that does not keep up with the "logic of prediction"

The Mahalanobis distance can also be used to set a criterion to detect such deviations.

The calculation of the properties uses 25 different regions of the spectrum. If the database uses N samples, then an unknown sample is acceptably expressed if the Mahalanobis distance D^2 is less than $(3 \times 25) / N$, according to the general formula:

$$D^2 < (3 \cdot M) / N \quad (M=\text{number of the spectrum regions, } N=\text{number of samples in the database})$$

If D^2 is greater than this limit, the unknown sample should be considered divergent (outlier), and the results on the properties of the samples showing this deviation will have, on average, lower accuracy.

If such divergent results are found regularly, the best solution is to construct a library with them if there are some properties or even one known property. Their input in the libraries is performed as previously described.

7.2.2 Change of the MLR model

This function described here is accessible for users of fifth or higher level.

As previously explained in the CA and the MLR section, Eraspec uses the MLR method when the CA can not be applied. For this calculation, regions of the spectra that match well with the calculated properties are selected.

A specific selection of spectral regions is called "model".

This selection of the model can be changed if the components present in the gasolines are not present in the standard MLR model, or vice versa.

	Time	Date	Sample-ID	RON	MON	Benz
1	17:41	20.12.12	gasoline	96.9	86.4	0.6
2	17:40	20.12.12	New solvent			
3	17:36	20.12.12	New solvent			
4	17:15	20.12.12	gasoline	95.9	85.4	0.6
5	16:29	20.12.12	gasoline	97.0	86.1	0.6
6	15:44	20.12.12	gasoline	94.9	86.7	0.2
7	15:43	20.12.12	gasoline	95.3	87.0	0.1
8	14:45	20.12.12	Test	95.6	87.2	0.1
9	14:42	20.12.12	Test	95.7	87.2	0.1
10	14:10	20.12.12	Test	95.8	87.2	0.2
11	14:03	20.12.12	Test	95.7	87.1	0.2
12	13:54	20.12.12	gasoline	95.4	87.0	0.2
13	13:46	20.12.12	gasoline	95.5	86.9	0.1
14	12:24	20.12.12	gasoline	94.9	86.5	0.1
15	12:15	20.12.12	gasoline	123.9	105.8	

In the menu of the results the Edit: Result key is pressed (*Fig.7.2.2.a*).

Fig.7.2.2.a

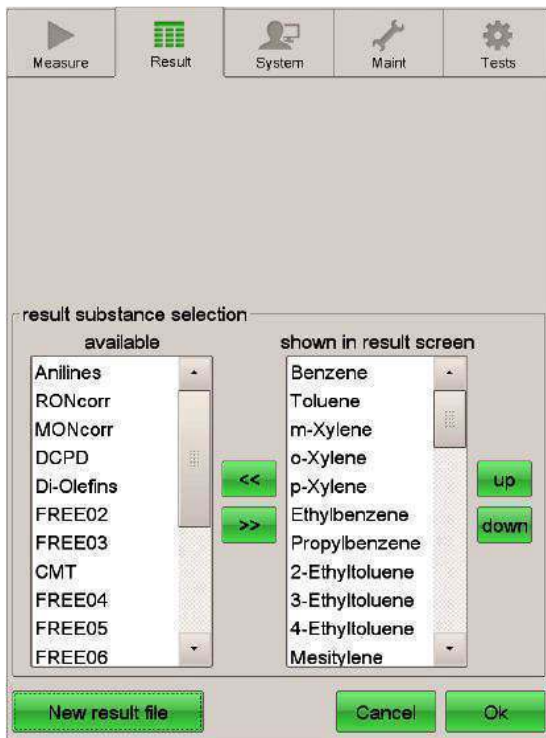


Fig.7.2.2.b

In the right white box, the user selects the parameter for which the model must be changed by touching the parameter name, and then selecting the submenu "Corr" in the top box (Fig.7.2.2.b).

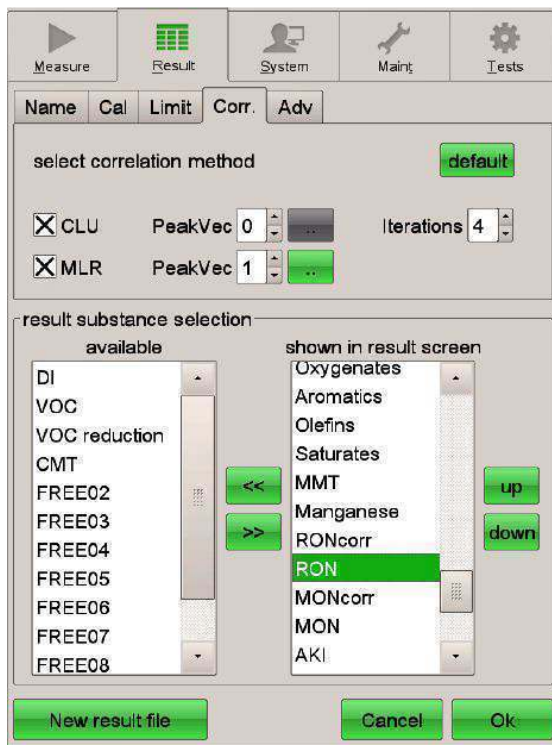


Fig.7.2.2.c

The method of correlation is selected by ticking the boxes CLU or/and MLR. The CLU is the CA (Cluster Analysis). Usually both methods are selected (Fig.7.2.2.c).

7.2.3 The statistical method of Eraspec

For the MLR method, there may be used some statistical functions to calculate how well the model can reproduce the known standard values entered in the database, how much linear is the model and if there are failure points (outliers) in the database.

The statistical quantitative relationships that Eraspec uses, if n is the number of samples in the library of known values, is:

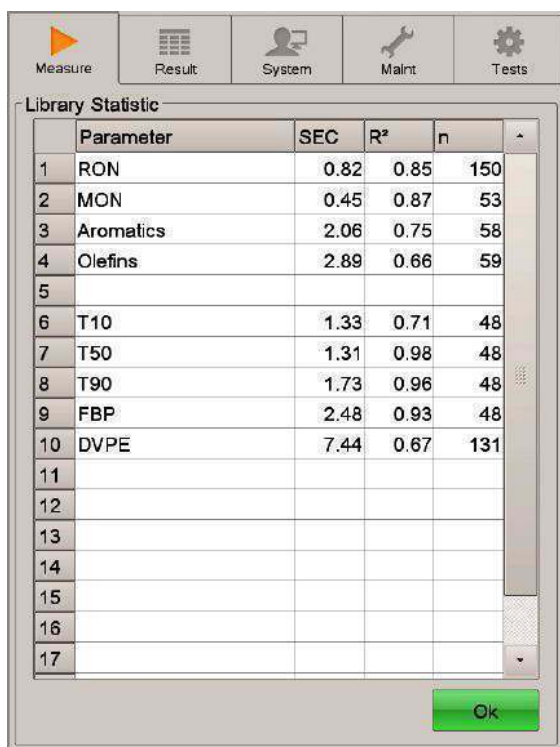
$$SEC = SQR \cdot [(P_{MLR} - P_{REF})^2 / n]$$

where: SEC (Standard Error of Calibration): is the standard deviation of the calculated values and the reference values for all n samples in the library.

Coefficient of determination, R^2 : this gives an estimate of how well a linear model can reproduce the known reference values. The R^2 is a perfect correlation and R^2 means no correlation at all.

Obviously these statistical values will be calculated only for the used libraries.

The following tables (Fig.7.2.3.a and 7.2.3.b) show the function of the reference values (x axis) with the values calculated by MLR (y axis):



Parameter	SEC	R ²	n
1 RON	0.82	0.85	150
2 MON	0.45	0.87	53
3 Aromatics	2.06	0.75	58
4 Olefins	2.89	0.66	59
5			
6 T10	1.33	0.71	48
7 T50	1.31	0.98	48
8 T90	1.73	0.96	48
9 FBP	2.48	0.93	48
10 DVPE	7.44	0.67	131
11			
12			
13			
14			
15			
16			
17			

Fig.7.2.3.a

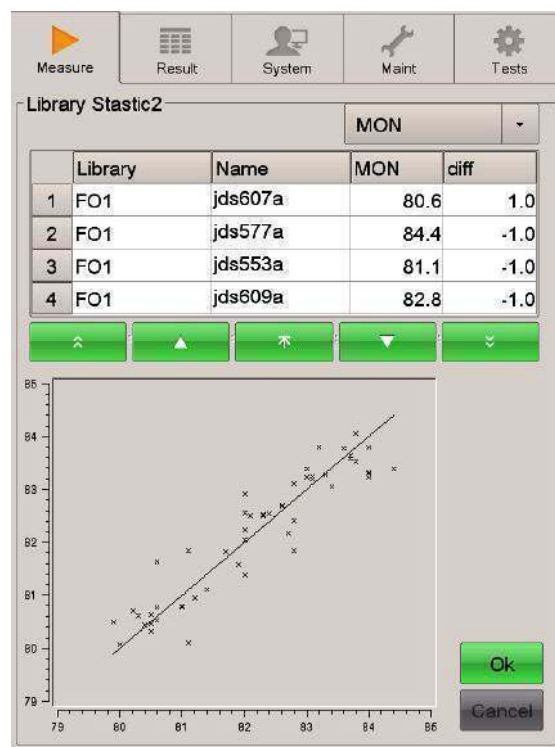


Fig.7.2.3.b

7.2.4 The selection of a different MLR model

In MLR line, the number displayed in a white box next to the "PeakVec" indicates the model used for calculating the MLR. There are twelve standard models for different types of gasoline and different properties available. If it is necessary the model for a selected parameter to be changed into a different model, the user is asked to enter the number of the desired model in the white box next to "PeakVec" in MLR line.

For each calculated parameter, the model can be changed this way, as shown below in the table Standard MLR models.

Press the key "OK" to save the changes and return to the results menu ("Result"). If "Cancel" is pressed, the changes that were made will not be saved, and the return to the menu "Result" occurs.

On the next page are shown in tabular form (*Table 7.2.4*) the standard MLR models of the instrument.

Table 7.2.4
Standard MLR models

Model Nr	Used for	Peak Areas used
0	IBP, T10, T50, T90 for gasolines without alkylate, with MTBE and/or Ethanol	Const, Density, P0078, P0111, P0170, P0194, P0239, P0263, P0286, P0305, P0447, P0487, P0583, P0682, P0695, P0733, P0797, P0803, P0999, P1303
1	RON, MON, FBP for gasolines without alkylate, with MTBE and/or Ethanol	Const, Density, P0078, P0111, P0170, P0194, P0239, P0263, P0286, P0305, P0447, P0487, P0583, P0733, P0797, P0803, P1303
2	Aromatics for all types of gasoline	Const, P0078, P0111, P0170, P0194, P0239, P0263, P0286, P0305, P170q
3	Olefins for all types of gasoline	Const, P0447, P0487, P0583
4	IBP, T10, T50, T90 for gasolines without alkylate, with MTBE and/or Ethanol and/or ETBE	Const, Density, P0078, P0111, P0170, P0194, P0239, P0263, P0286, P0305, P0447, P0487, P0583, P0682, P0695, P0733, P0777, P0797, P0803, P0999, P1303
5	RON, MON, FBP for gasolines without alkylate, with MTBE and/or Ethanol and/or ETBE	Const, Density, P0078, P0111, P0170, P0194, P0239, P0263, P0286, P0305, P0447, P0487, P0583, P0733, P0777, P0797, P0803, P1303
6	IBP, T10, T50, T90 for gasolines with alkylate, without oxygenates	Density, P0078, P0111, P0170, P0194, P0239, P0263, P0286, P0305, P0447, P0487, P0583, P0682, P0695, P1303, P0856, P0939
7	RON, MON, FBP for gasolines with alkylate, without oxygenates	Const, Density, P0078, P0111, P0170, P0194, P0239, P0263, P0286, P0305, P0447, P0487, P0583, P1303, P0856, P0939
8	IBP, T10, T50, T90 for gasolines with alkylate, with MTBE and/or Ethanol and/or ETBE	Density, P0078, P0111, P0170, P0194, P0239, P0263, P0286, P0305, P0447, P0487, P0583, P0682, P0695, P0733, P0777, P0797, P0803, P0999, P1303, P0856, P0939
9	RON, MON, FBP for gasolines with alkylate, with MTBE and/or Ethanol and/or ETBE	Const, Density, P0078, P0111, P0170, P0194, P0239, P0263, P0286, P0305, P0447, P0487, P0583, P0733, P0777, P0797, P0803, P1303, P0856, P0939
10	IBP, T10, T50, T90, FBP for gasolines without alkylate, with MTBE	Const, Density, P0078, P0111, P0170, P0194, P0239, P0263, P0286, P0305, P0447, P0487, P0583, P0682, P0695, P0797, P1303, P170q
11	RON, MON, Distillation for gasolines with alkylate, with MTBE and/or Ethanol and/or TAME	Const, Density, P0078, P0111, P0170, P0194, P0239, P0263, P0286, P0305, P0447, P0487, P0583, P0682, P0695, P0733, P0797, P0976, P1303, P170q, P0856, P0939
12	IBP, T10, T50, T90 for gasolines with alkylate, with MTBE and/or Ethanol	Density, P0078, P0111, P0170, P0194, P0239, P0263, P0286, P0305, P0447, P0487, P0583, P0682, P0695,
		P0733, P0797, P0803, P1303, P0856, P1075
13	RON, MON and FBP for gasolines with alkylate, with MTBE and/or Ethanol	Density, P0078, P0111, P0170, P0194, P0239, P0263, P0286, P0305, P0447, P0487, P0583, P0733, P0797, P0803, P1303, P0856, P1075

7.2.5 Change and creation of a model

The standard models are specific for gasoline types for which a factory default database is available.

It may be desirable to create or change a new model, for example if a new oxygenate is present in gasolines for which currently there is no database and therefore there is not any ready model available.

In the results menu, press the "Result" key and choose a calculated property. Then click the ".." button next to "PeakVec":

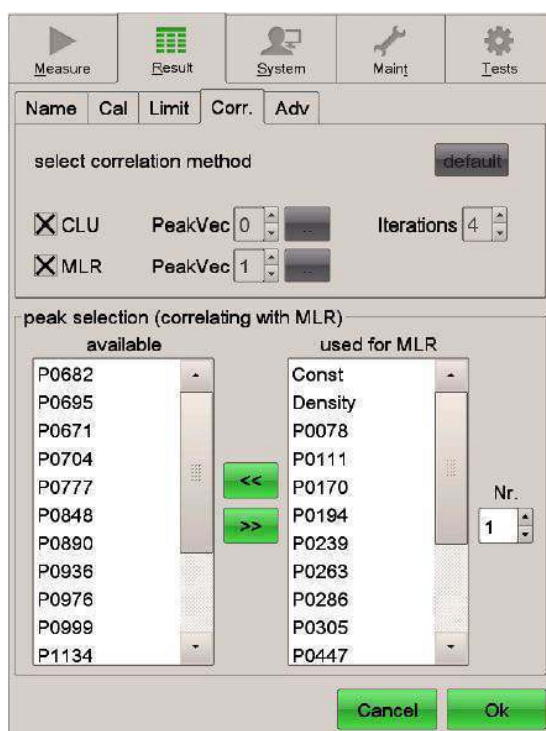


Fig.7.2.5.a

In the white box below the "Nr" select either the number of the existing standard model to be changed (previous section "Standard MLR models"), or a number between 12 and 20 to create a new model.

The white box "used for MLR", shows the regions of the peaks that are defined for the selected model.

In the white box "available", all the available additional peaks are depicted.

To add a peak in the model, the user should touch its name in the box "available". The selection will turn green. Then the user has to move this peak into the box "used for MLR" by touching the button ">>".

The process is repeated so that all the peaks are moved to the selected model.

To remove a peak is the reverse process i.e. from the column "Used for MLR" move the peak to the column "Available".

To save the changes press the "OK" key. If "Cancel" is pressed the changes are not saved.

To use a new model for a calculated property, the correct model number must be entered.

Select the name of the calculated property from the list in the box "shown in result screen" (Fig.7.2.5.b).

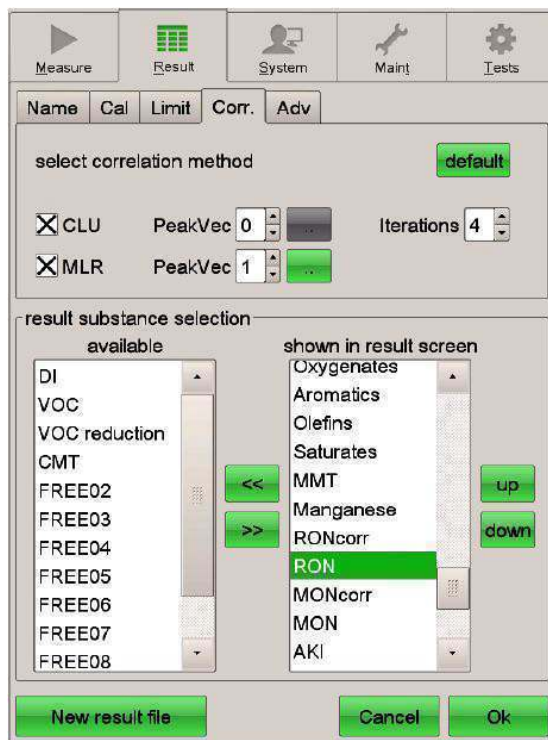


Fig.7.2.5.b

Then the user changes the number in the white box next to "PeakVec" to the number of the desired model.

Repeat this process for all the calculated properties for which a new model must be used.

Respectively press "OK" to save and "Cancel" to cancel.

7.2.6 The usefulness of FREE parameters

It may be desirable to calculate other parameters of the infrared spectrum as well. Therefore, apart from the standard parameters that exist when the user enters data in the library, there are also ten parameters which are called FREE and are defined by the user.

These are FREE01, FREE02, FREE03 FREE10.

An example is a distillation property e.g. "Evaporate at xxx", where xxx is some specific temperature. This temperature in some countries is expressed in Fahrenheit scale while in others in Celsius scale. The specifications as well, may vary from country to country.

Eraspec has 10 additional such parameters which may be renamed and used in the database to make it possible to calculate the additional parameters. Of course, at least 40 samples with these properties being known must be entered in the libraries in order for them to contribute to the calculation of an unknown sample later.

During the calibration process the text editor renamed the FREE01 into "Sulfur" (sulphur) measured in "ppm" and the parameter FREE02 into "DVPE" and in the check box for entering the units he selected the option "unit". In other words, for Sulfur in the white box he introduced ppm and for DVPE, because unit was not needed, he introduced the word "unit".

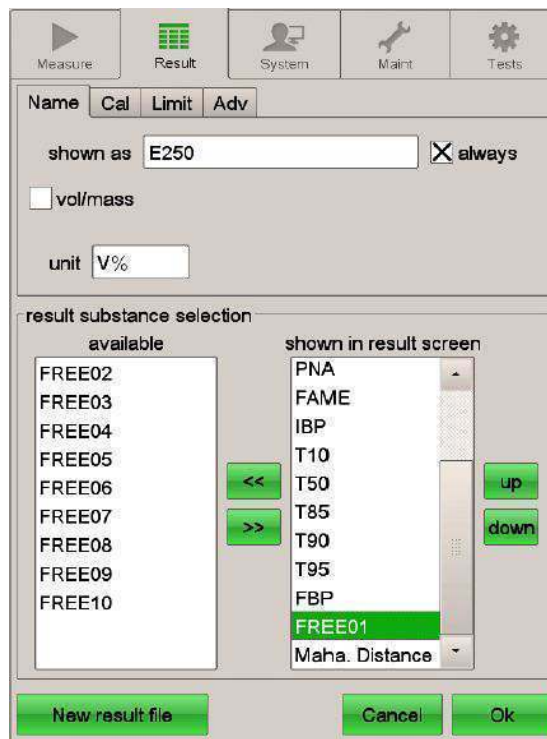


Fig.7.2.6

After entering more than 40 samples it was observed that this option was working in unknown samples, but with huge deviations. For example, while sulphur in the certificates was from 7.8 to 10 ppm with a maximum of 11, in the unknown samples it appeared with values 67 to 70 ppm.

After contacting the manufacturer's technician it was understood that:

One can only affect the parameters calculated by the database by adding more samples in it and by setting the values. This however does not include properties such as benzene for example. These items are separated by other means and are internationally calibrated.

FREE parameters can be used to train the instrument to calculate brand new parameters using new signals (which are called "peak vectors") from the spectrum, which are already set for other parameters. However this does not include new substances. If it is desired for example, an alcohol, pentanol, to be measured, the spectrum should be measured in the factory and then introduce it in the instrument's software. Only then the instrument will be able to measure the concentration of pentanol in the sample. Of course this works, if the new substance can be detected by the instrument and the spectrum of the substance (the peak corresponding to it) can be distinguished from other substances' peaks.

Below is shown a list of all the peaks vectors in the software and can be used for the MLR to work. Addition of new vector peaks is only possible in the factory and in no case this can be done by the user.

On the next page there is a table (*Table 7.2.6*) containing all the peaks and their corresponding substances. [13]

Table 7.2.6
Peaks and the corresponding models

Peak caused by	Type	Peak in PeakVec
Benzene	Aromatic	P0078
Toluene et al	Aromatic	P0111
Toluene	Aromatic	P0170
o-Xylene	Aromatic	P0194
m-Xylene	Aromatic	P0239
Naphthalene	Aromatic	P0263
p-Xylene	Aromatic	P0286
Pseudocumol	Aromatic	P0305
Olefins	Olefin	P0447
Olefins	Olefin	P0487
Olefins	Olefin	P0583
Aromatics	Aromatic	P0682
Aromatics	Aromatic	P0695
Methanol	Oxygenate	P0704
Ethanol	Oxygenate	P0733
ETBE	Oxygenate	P0777
MTBE	Oxygenate	P0797
Ethanol	Oxygenate	P0803
ETBE	Oxygenate	P0848
Dimethoxymethane	Oxygenate	P0890
DIPE	Oxygenate	P0936
TAME, tert-Butanol	Oxygenate	P0976
MTBE	Oxygenate	P0999
Dimethylcarbonate	Oxygenate	P1134
Methyl 1377 cm ⁻¹	-	P1303
Aromatics	Aromatic	P1702
Toluene^2	Aromatic	P170q
Alkylate	Alkylate	P0856
Alkylate	Alkylate	P0939
Alkylate	Alkylate	P1075

8. Exporting spectra from the device

After enrichment of the libraries, some measurements of samples from intermediate distillation products were conducted, in order for them to be compared to each other, but also with gasoline spectrum from gas station.

8.1 Identification techniques

Before the citation of spectra and the comparison between them, it is important to develop a theoretical part which will include a simple strategic for manual identification of spectra. There are some general rules that can be used when an infrared spectrum is studied for the identification of a molecular structure. These rules are as follows:

- 1) First the observer looks at the largest wavenumber at the end of the graph ($>1500\text{ cm}^{-1}$) and initially is concentrated at the large peaks.
- 2) For each peak a short list is constructed, using a correlation chart.
- 3) In the start of the spectrum the lower wavenumber is used for the confirmation, or for the treatment of possible structural elements.
- 4) Apparently the observer should not expect to be able to identify each peak in the spectrum.
- 5) There should be continuous monitoring for "crossing" with a possible substance anywhere possible. For example, an aldehyde must absorb near 1730 cm^{-1} and in the region $2700\text{-}2900\text{ cm}^{-1}$. These were analyzed and are found extensively in Chapter 4 of the theory.
- 6) The observer must take advantage of both positive and negative signs of the spectrum. For example if there is no absorption in the region $1600\text{-}1850\text{ cm}^{-1}$, then more than likely there is no carbonyl group in the molecule.
- 7) Specific intensities of the lines should be encountered carefully. Under certain conditions, they might vary considerably even for the same group.
- 8) Caution should be given when small changes in wavenumbers are studied. These can be affected by the fact that a spectrum is obtained from solid or liquid, or in a solution. If extracted from solution, some peaks are very "solution-sensitive".
- 9) We should not forget to subtract the solvent's spectrum if possible, because it would be confused with that of the sample.

To make a clear identification through the examination of the infrared spectrum of a complex component such as fuel it is not always possible.

In older times, the identification was made manually with atlases containing too many spectra. As understood this process was very tedious and time consuming. When the substance was identified, it was added in the existing atlases with its turn and enriched them.

8.2 Sample types

For the observation of differences between the spectra of fractions of gasoline, seven spectra from seven intermediate distillation stages were taken. More specifically the spectra of the following fractions were obtained:

- 1) ETBE
- 2) FCC C6 +
- 3) Alkylate
- 4) C5 Raffinate + TAME
- 5) Reformate
- 6) Isomerate
- 7) LSR
- 8) Gas station gasoline 95 R

Regarding diesel the spectra from five different samples were obtained:

- 1) Light Cycle Oil, MOH
- 2) Hydrocracker, Eleusis, HELPE
- 3) Hydrotreater, Aspropyrgos, MOH
- 4) Hydrotreater Straight run, Eleusis, HELPE
- 5) HVO, HELPE
- 6) Motor Diesel from gas station
- 7) Diesel for heating

8.2.1 Sample processing

The export of spectra with USB is itself a special process, but in this work there is no meaning for its analysis. More meaningful is the data processing with Excel because there the values are pure and errors and discrepancies are seen clearly. In particular, the instrument after each measurement gave a file after every 5900 values corresponding to absorption values. The absorbance is matched to a wavelength. Based on the instructions book, the first absorption value is the wavenumber 631 up to number 4100, with steps of 0.5736.

The wavenumber is converted to wave length by the relation:

$$k = 2 \cdot \pi / \lambda$$

If k is converted to nanometers the device works from 9.968 to 1570 nm.

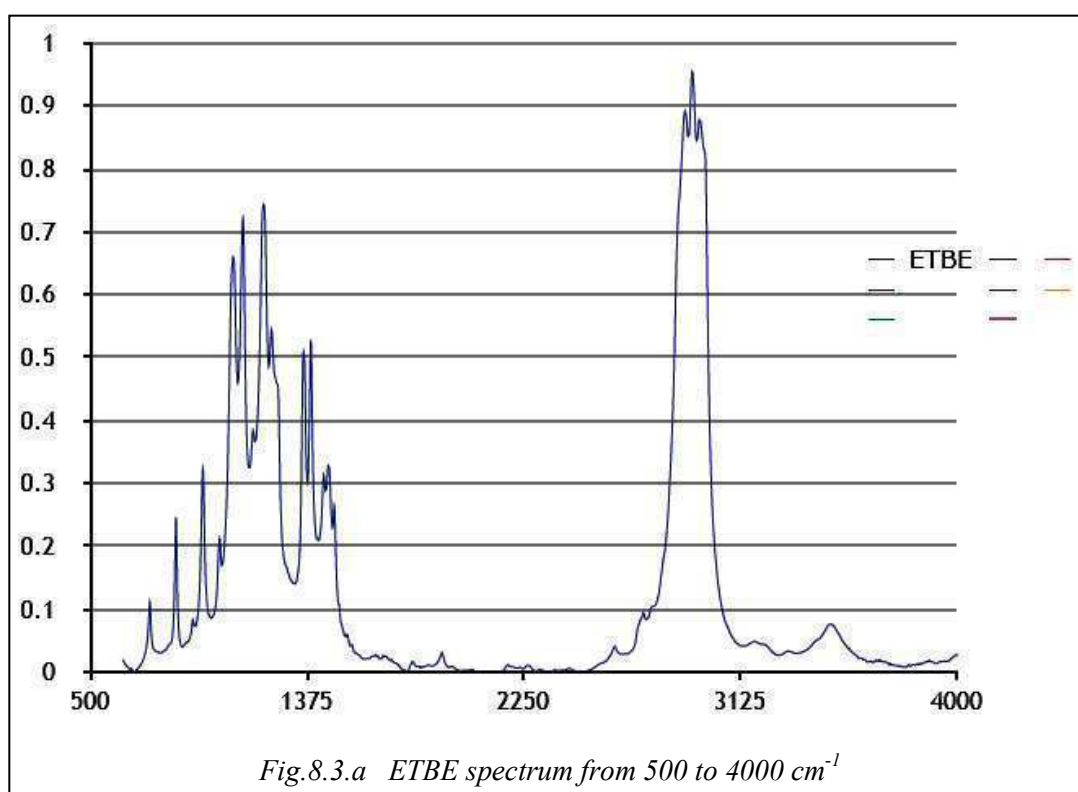
It is known that the absorption takes values from 0 to 1. In many cases, however, there were negative and also values greater than 1 resulted. It should be recalled that absorption is a dimensionless value. At 0 value no absorption occurs while at 1 it takes place fully. The continuous spectrum which is suitable for comparison and identification for most of these fractions is intermediate of 631 and 2000 cm^{-1} . [14]

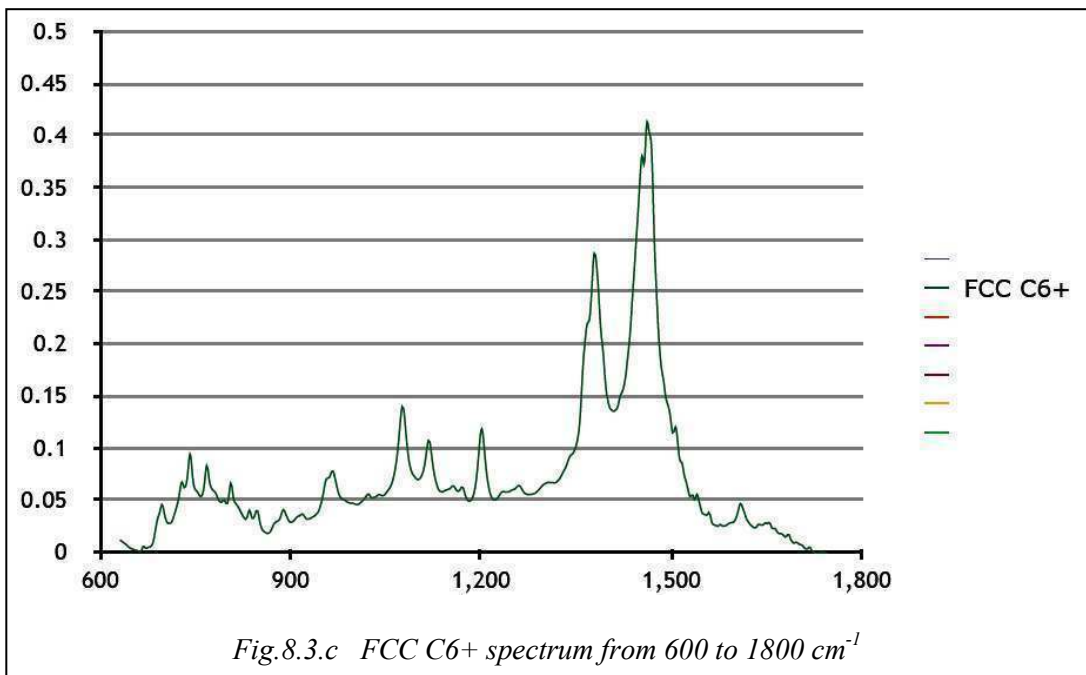
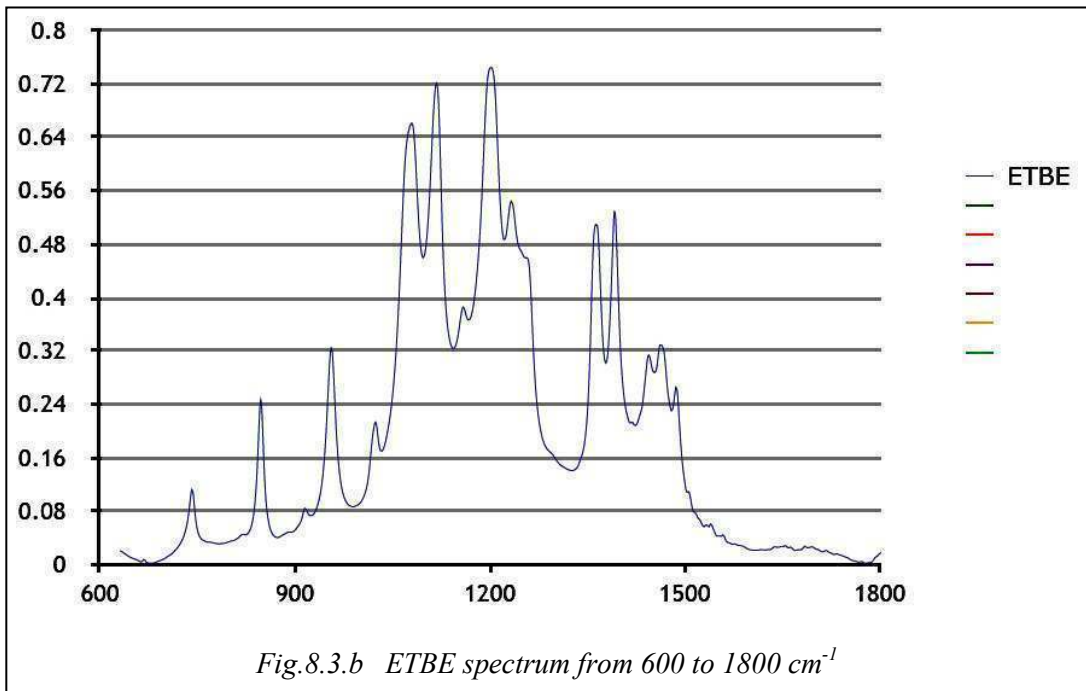
These errors are due to many factors. Perhaps the most common is the contamination of the sample by impurities. (Some of the samples were quite old). Errors due to the interferometer

also took place. If the mirror moves just a little slower than normally, the spectrum becomes linear for specific values. These are seen most clearly in the next chapter for spectra display.

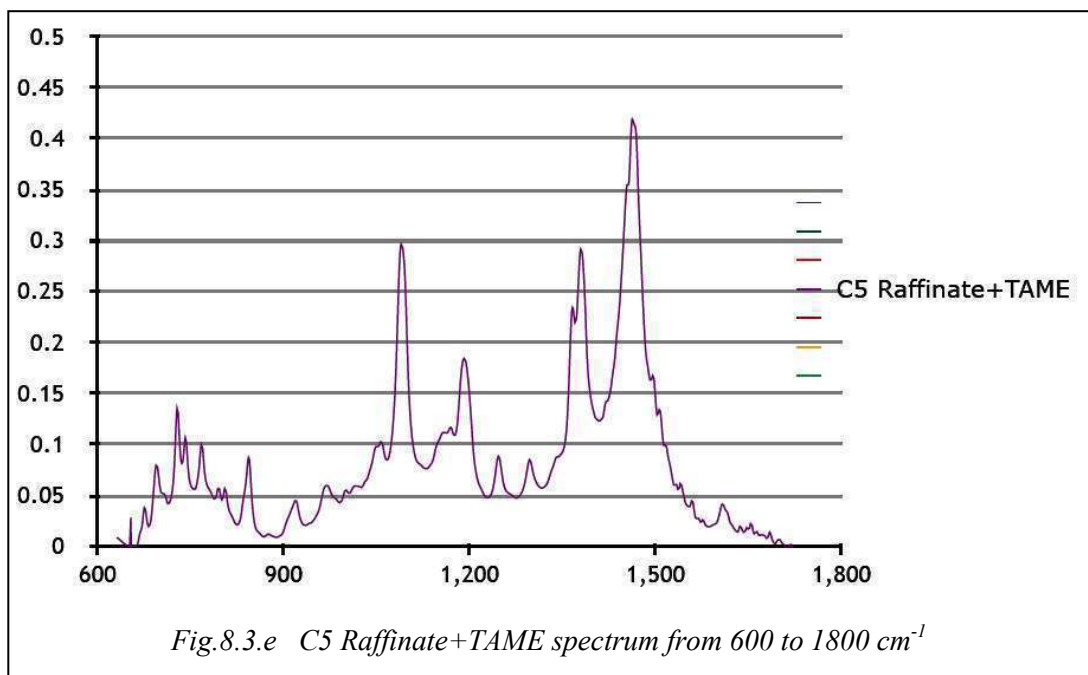
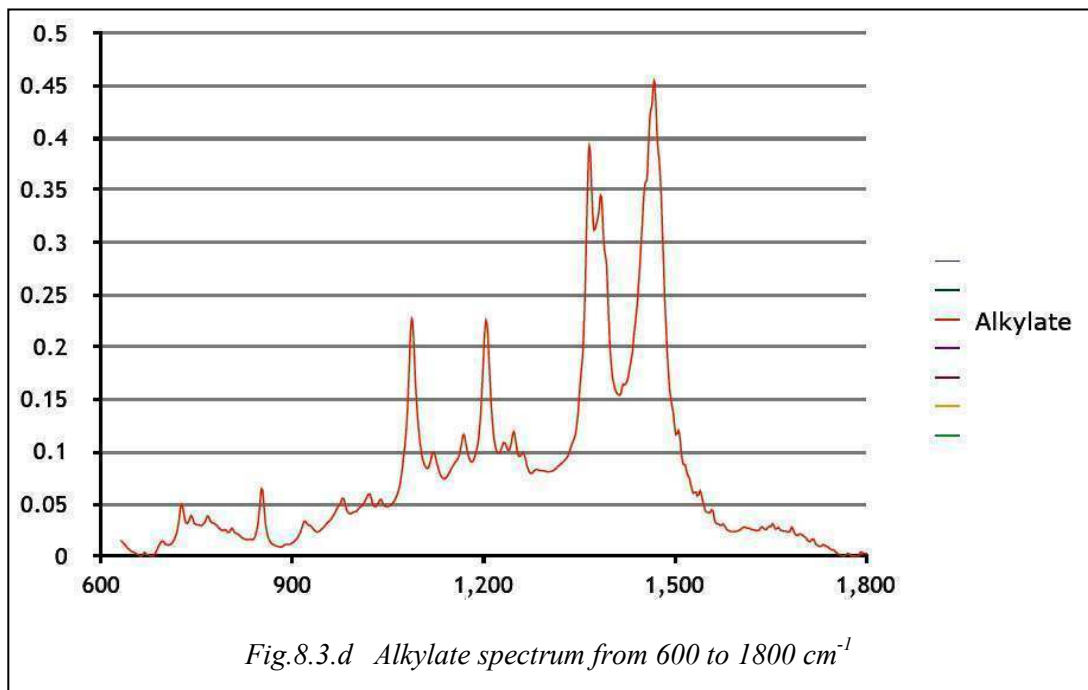
8.3 Display of spectra

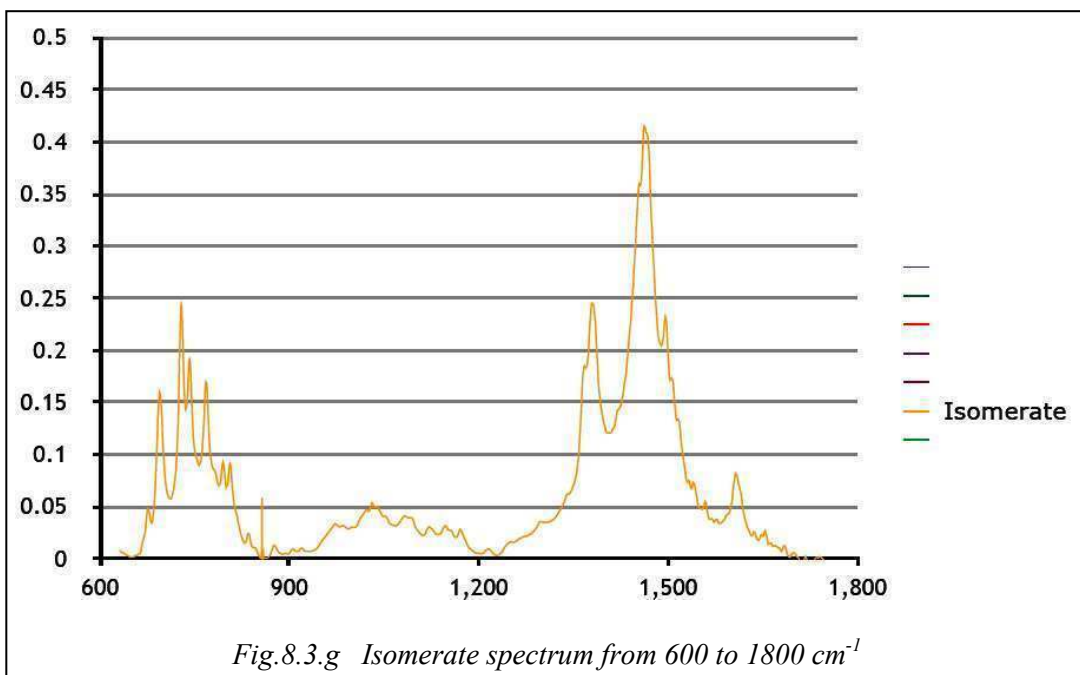
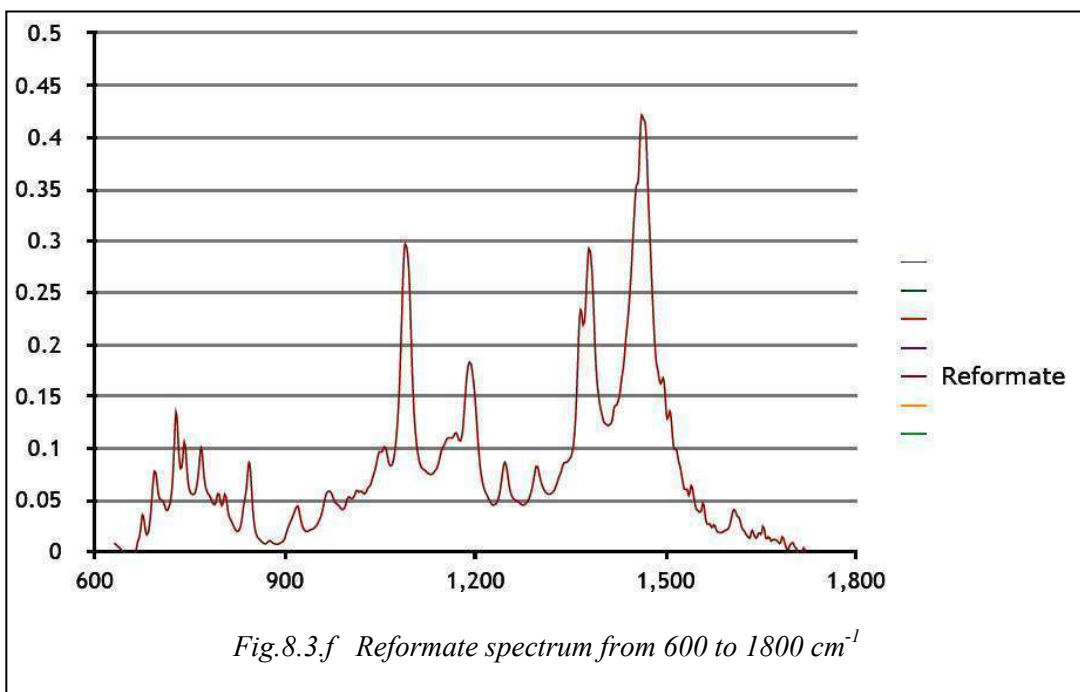
In Excel initially the spectra were made at a wavenumber from 631 to 4000 cm^{-1} at absorbance 0 to 1. However, since the functional region for comparison is from 631-1800 cm^{-1} , the spectra were converted to 631-1800 cm^{-1} and absorption from 0 to 0.5. This was happening because the high peaks absorbing at 3000 cm^{-1} , in most cases surpass the value of 1. In the functional region mentioned before, the maximum absorption is about 0.5.

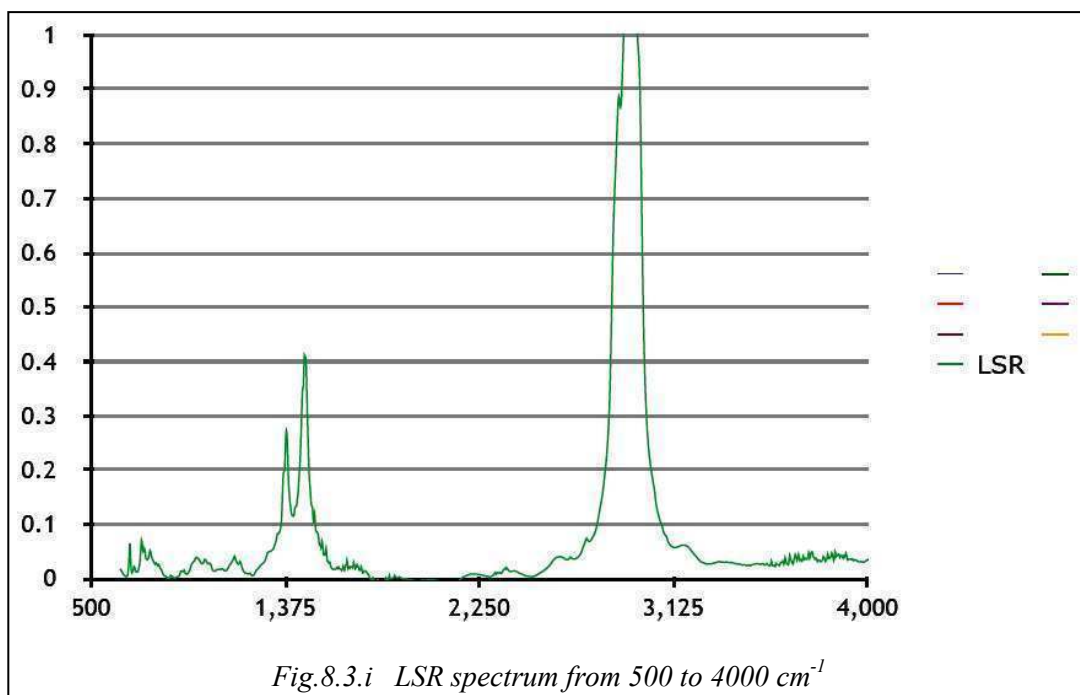
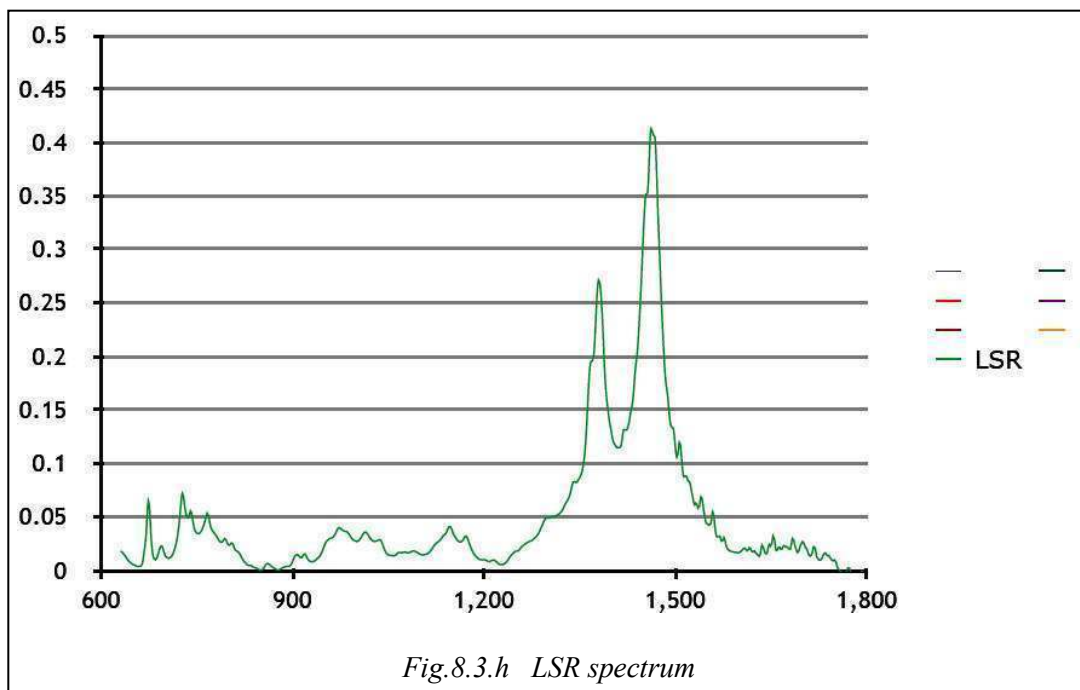




From this point on, regarding the presentation of the spectra, only the functional region of the spectrum will be displayed.

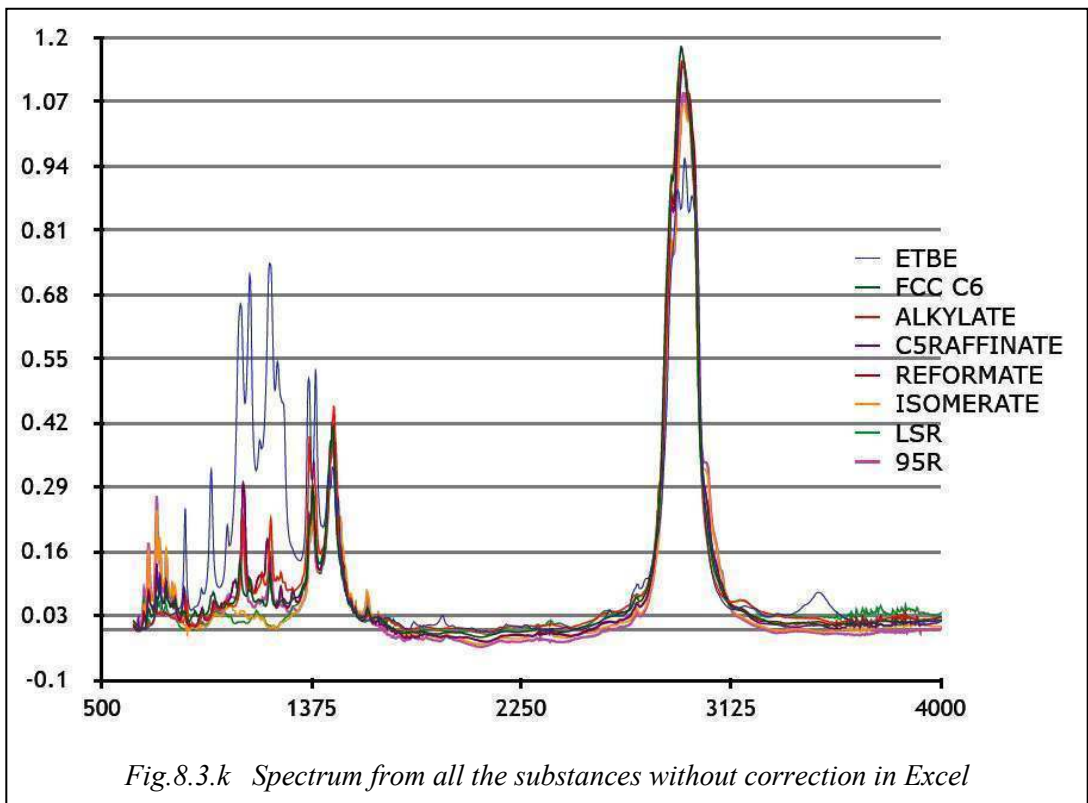
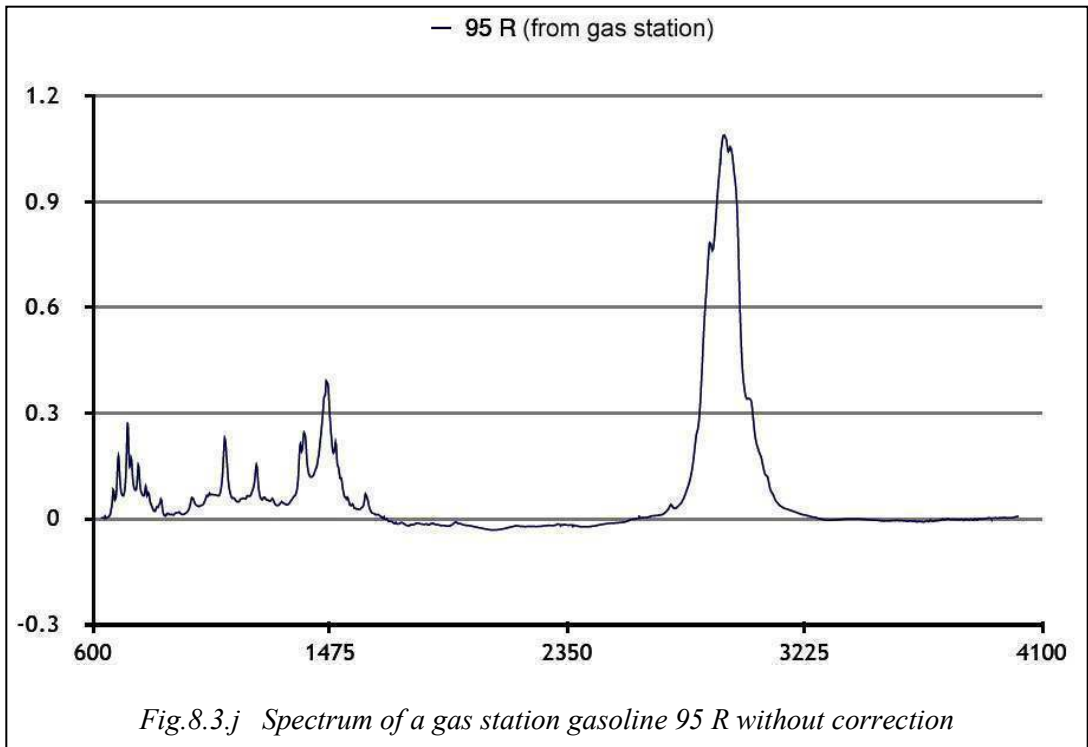






As seen above, the strong absorption at about 3000 cm^{-1} has exceeded 1. This is wrong and may be due to incorrect conversion of the analog signal to digital one.

On the next page the gasoline spectra joined all together in a common graph are displayed, so that their similarities and differences as to a spectrum of a regular gasoline 95 to be easily observed.



8.4 Comparison of gasoline spectra

In the whole spectrum it is observed, that among these substances there are big similarities. However significant differentiation is shown by ETBE. Moreover it is observed that all these fractions show a very intense absorbance in common, at about 2926 cm^{-1} . It is so intense that in most of the substances analyzed it seems as an error, since it surpasses 1. On the basis of the theory of Chapter 4, this "joint peak" is due to the stretching of the asymmetric C-H bonds in methylene group.

The rest of the comparisons that follow will be made after the isolation of the peaks of the spectra under comparison. This way the peaks will be more easily visible and more distinct.

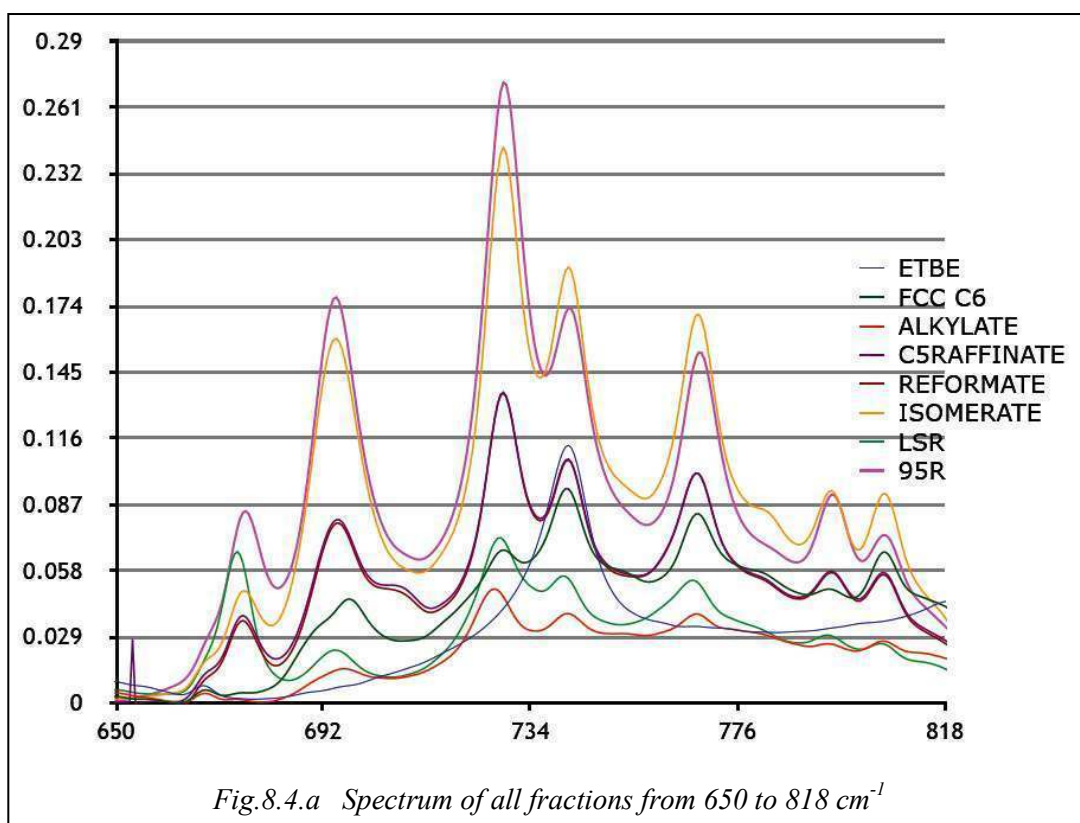


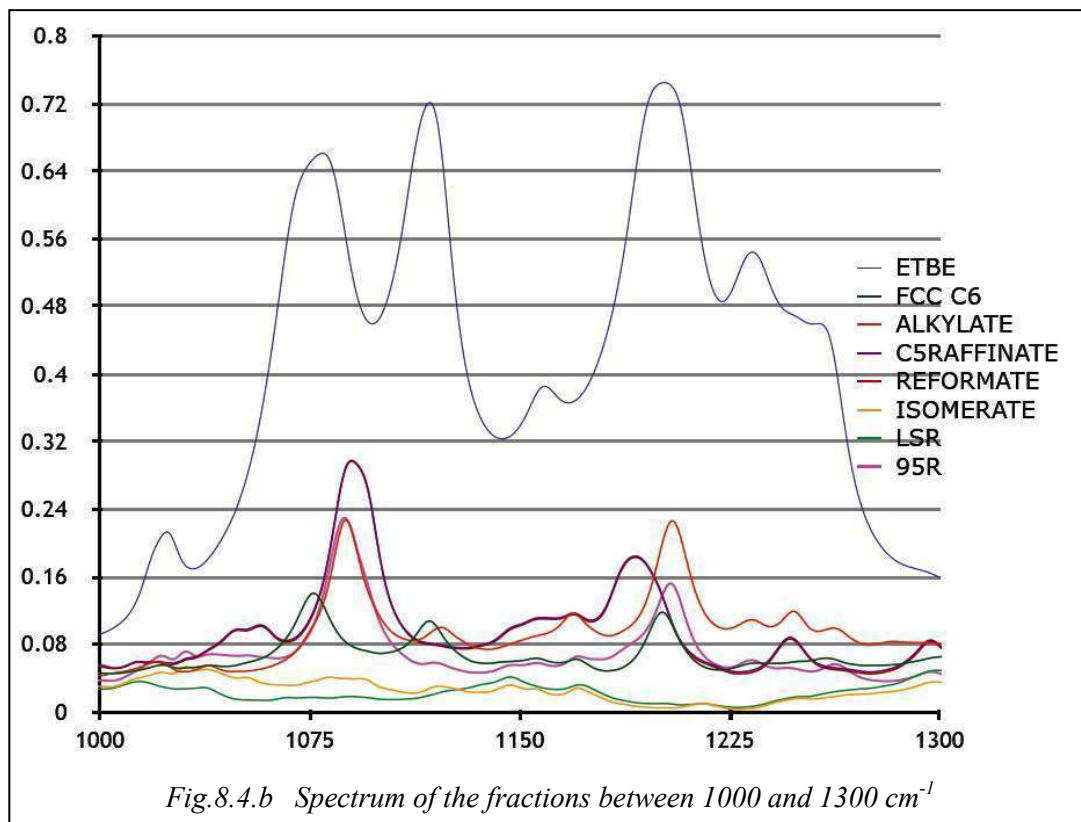
Fig.8.4.a Spectrum of all fractions from 650 to 818 cm^{-1}

This specific spectrum is from 650 to 818 cm^{-1} with absorbance values from 0 to 0.29 (Fig.8.4.a). Touching the mouse cursor onto the chart in Excel, the wavenumber where the absorptions occur, can be taken.

In this figure we see that there is strong absorption in the points with wavenumbers 676 , 695 , 728 , 742 , 768 , 795 and 806 cm^{-1} .

The greater absorptions are due to the higher concentration of the ingredients giving the peaks.

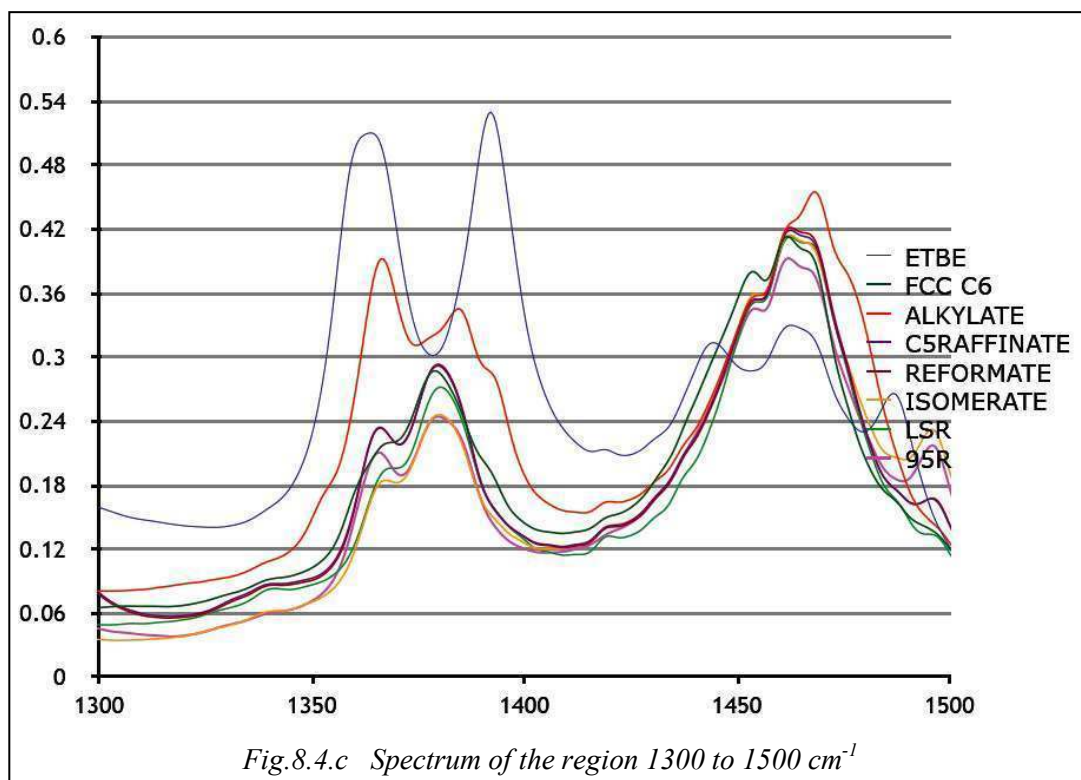
The correlations for the above wavenumbers are as follows: $=\text{C-H}$ bending, C-H bending, Methylene rocking, Double 1,2 aromatic, Double 1,2 aromatic, Triple 1,2,3 aromatic and Triple 1,2,3 aromatic. Generally the aromatics are from 740 up to 806 cm^{-1} .



In the above image (*Fig.8.4.b*) the spectrum between 1000 and 1300 cm⁻¹ was isolated. Strong absorptions are observed at points and regions: 1080-1095 cm⁻¹, 1119 cm⁻¹, 1190-1203cm⁻¹, 1234-1250 cm⁻¹.

Referring to the theoretical part, the conclusions respectively for the parts of the spectrum are as follows:

It is likely that all are due to in-plane C-H bending, because the peaks in the spectrum are not precisely defined but have indistinct boundaries. An exception may be at 1119 cm⁻¹ which is possibly due to SO₂ symmetric stretching bonding. It is observed, however, that ETBE at this specific region shows much more intense absorptions than the other fractions. Isomerate and LSR do not almost show any absorption at this specific region of wavenumbers. In addition there are many oxygenates contained, since intense absorptions are generally observed.



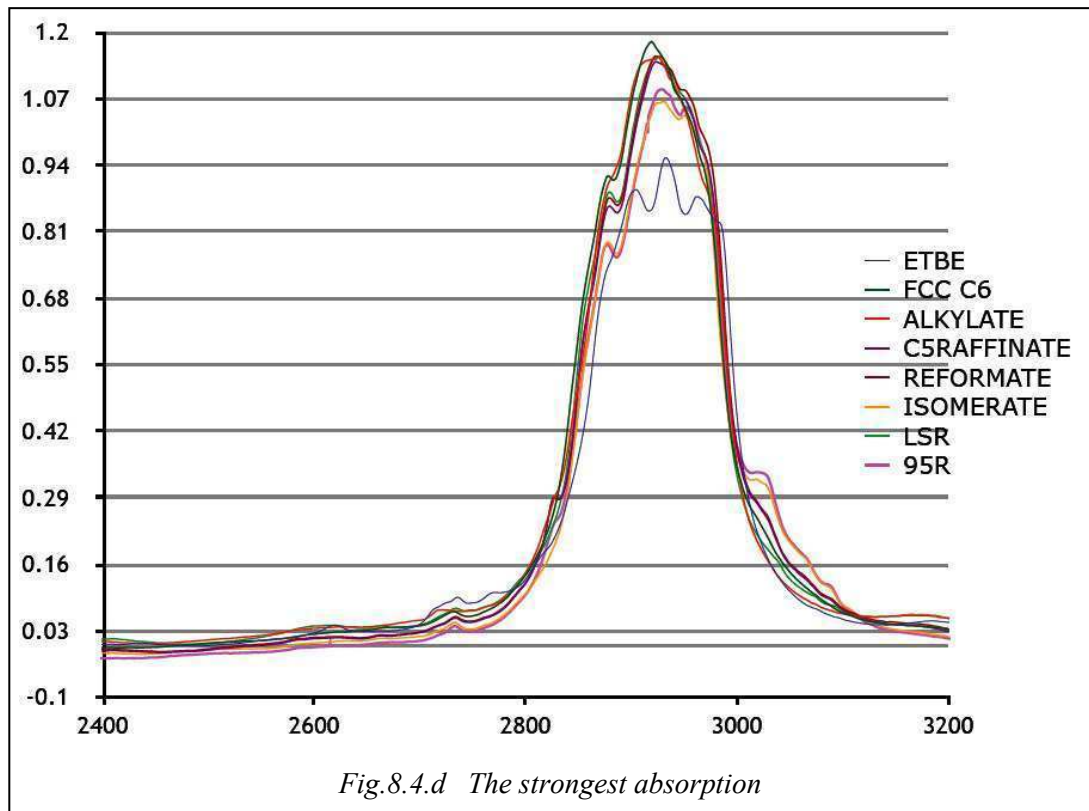
The characteristic peaks observed here (*Fig.8.4.c*) are at the regions: 1364-1365 cm⁻¹, 1380-1392 cm⁻¹, 1444-1455 cm⁻¹, 1463-1469 cm⁻¹, 1487-1497 cm⁻¹.

The correlations for the above wavenumbers are roughly as follows:

Nitrogenous aromatic multicomponents with NO₂, stretching bond, SO₂ symmetric stretching bond or Methyl symmetric C-H bending, double bond N=N, methylene group, very typical scissoring, C=C stretching bond.

Of particular interest is the study of the very large peak formed in all eight samples (*Fig.8.4.d*). This peak appears at wavenumbers 2924-2935 cm⁻¹ and absorbances close to 1, that is total absorption is observed. As mentioned in §2.3.3, with almost total absorption little light reaches the detector resulting in errors. Generally, in large or respectively in very faint absorptions the results are not accurate. This is due to signal-to-noise ratio.

It is observed that the stronger absorption appears at 2924-2935 cm⁻¹, while a smaller peak occurs before this, at approximately 2892 cm⁻¹. Both these peaks are due to aliphatic hydrocarbons, with the highest to Methylene asymmetric C-H stretching. The other is due to Methyl asymmetric C-H stretching.



8.5 Presentation of diesel spectra

The comparison between the spectra regarding diesel will be made exactly the same way as was done in the previous chapter. The only difference will be that all spectra will be presented in order from 631 to 2000 cm^{-1} i.e. in their consecutive functional area. Moreover in the case of diesel (as in gasoline previously) a very strong absorption is observed. This absorption will be explained in detail in this section as well.

Initially, the spectra are presented successively individually below.

Fig.8.5.a: LCO

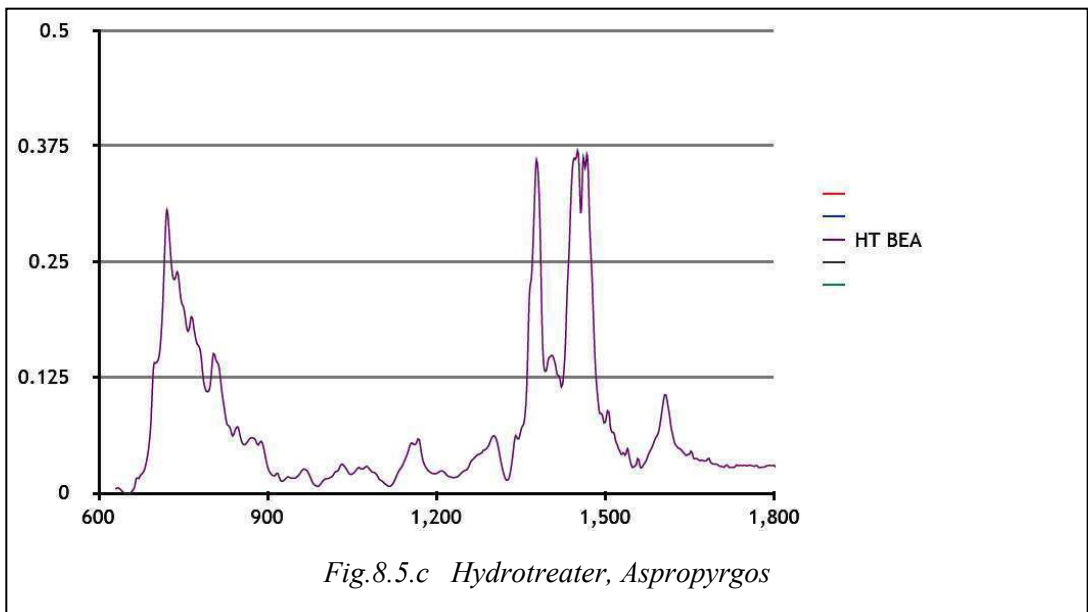
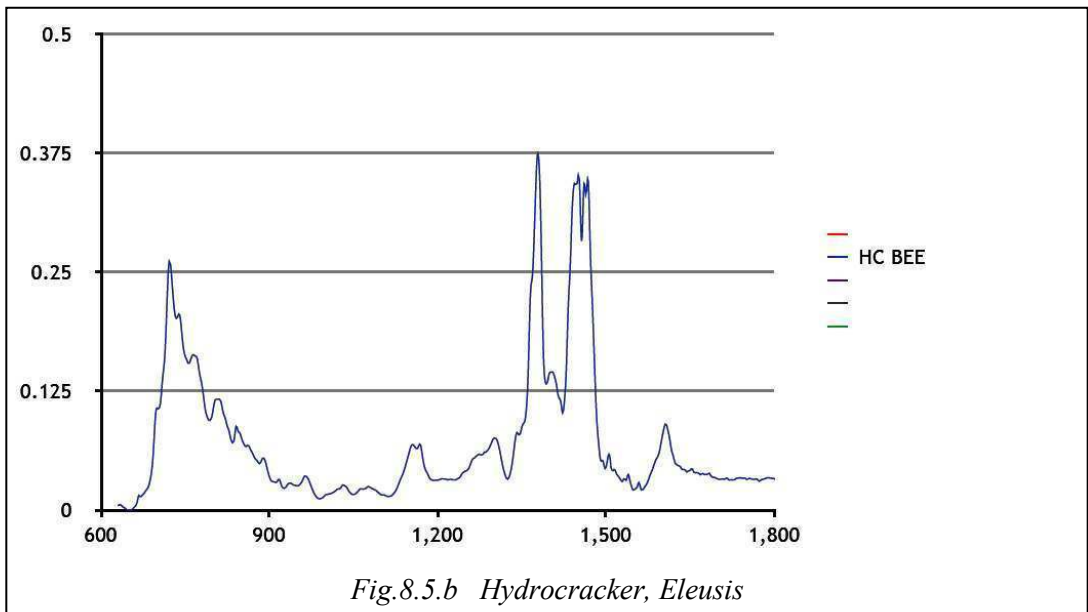
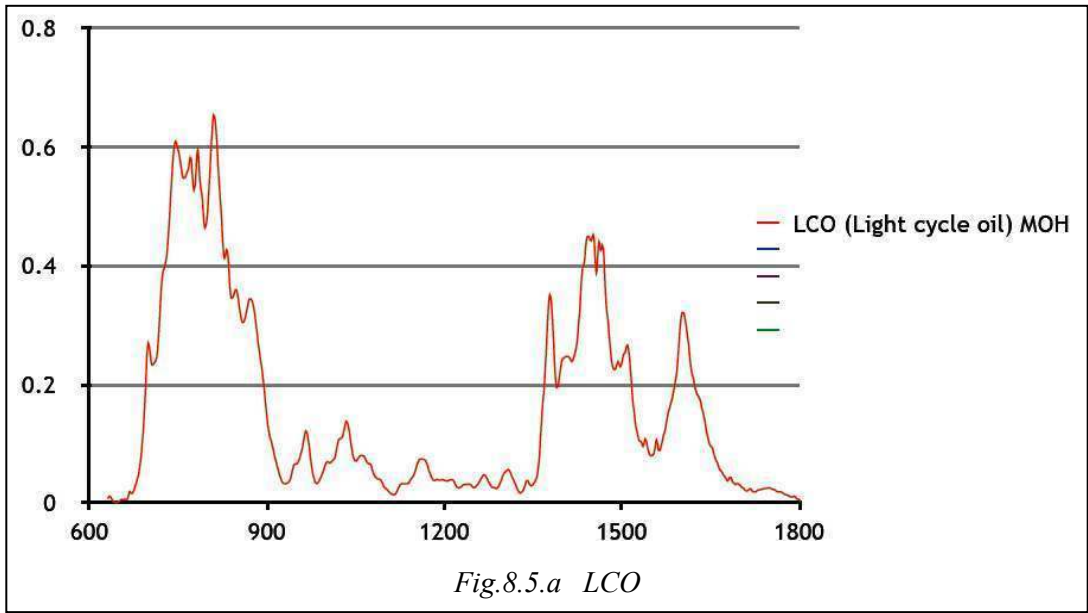
Fig.8.5.b: Hydrocracker, Eleusis

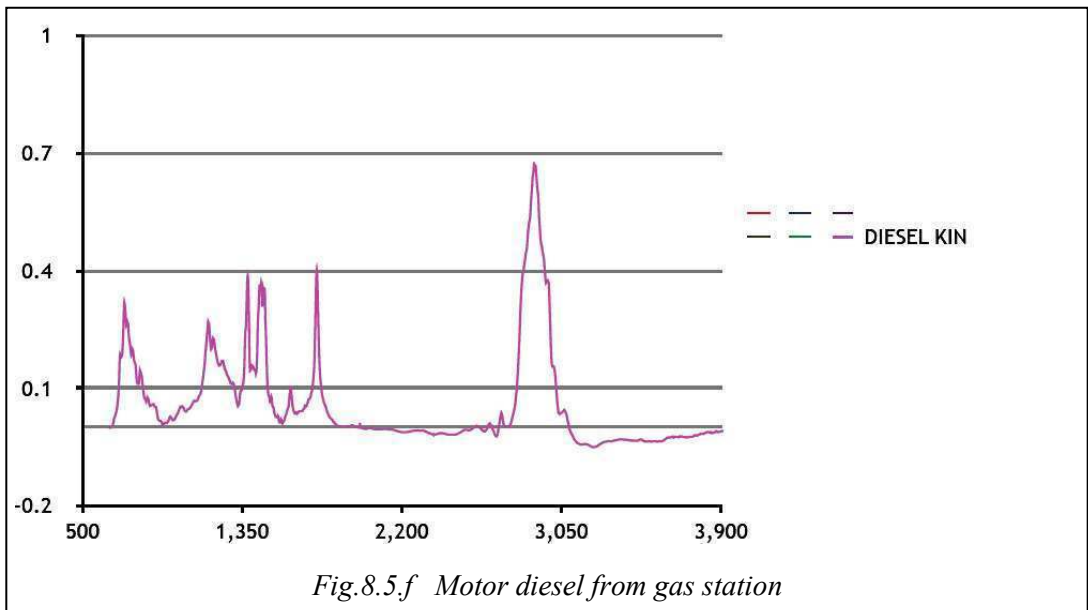
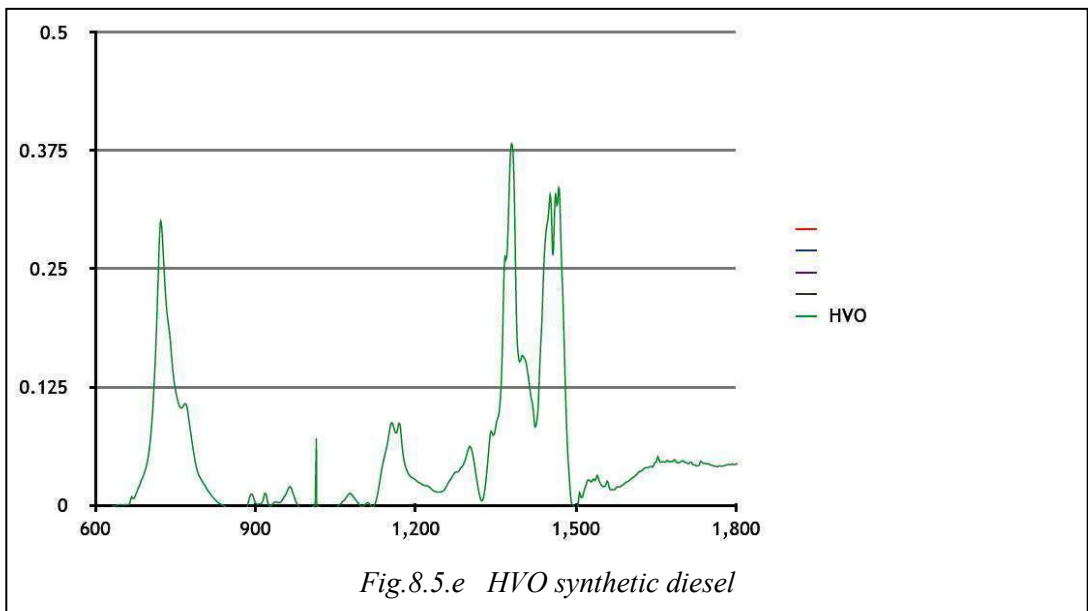
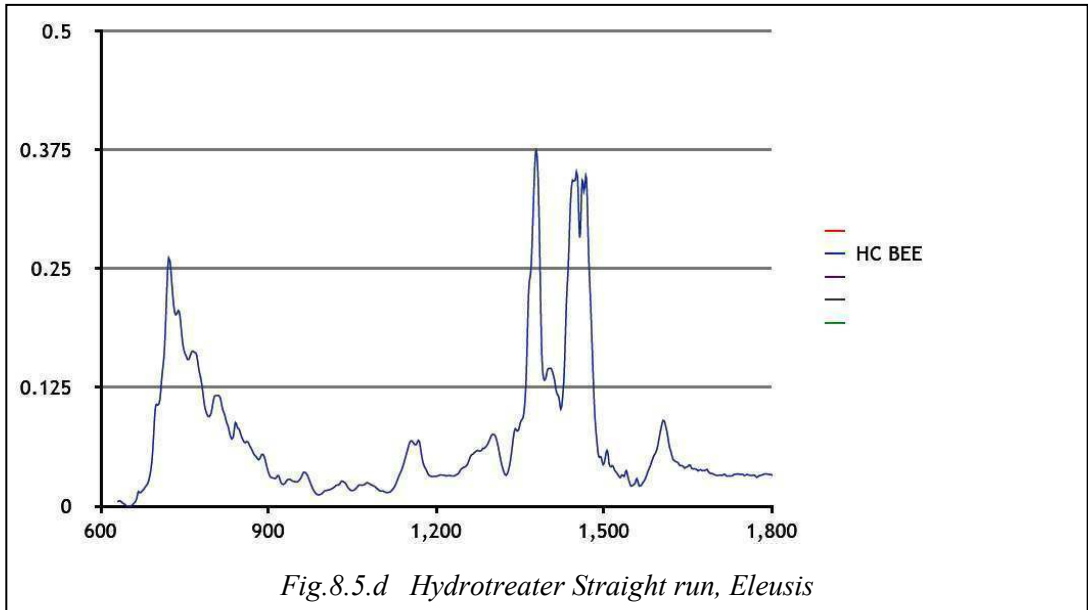
Fig.8.5.c: Hydrotreater, Aspropyrgos

Fig.8.5.d: Hydrotreater Straight run, Eleusis

Fig.8.5.e: HVO

Fig.8.5.f: Motor Diesel from gas station

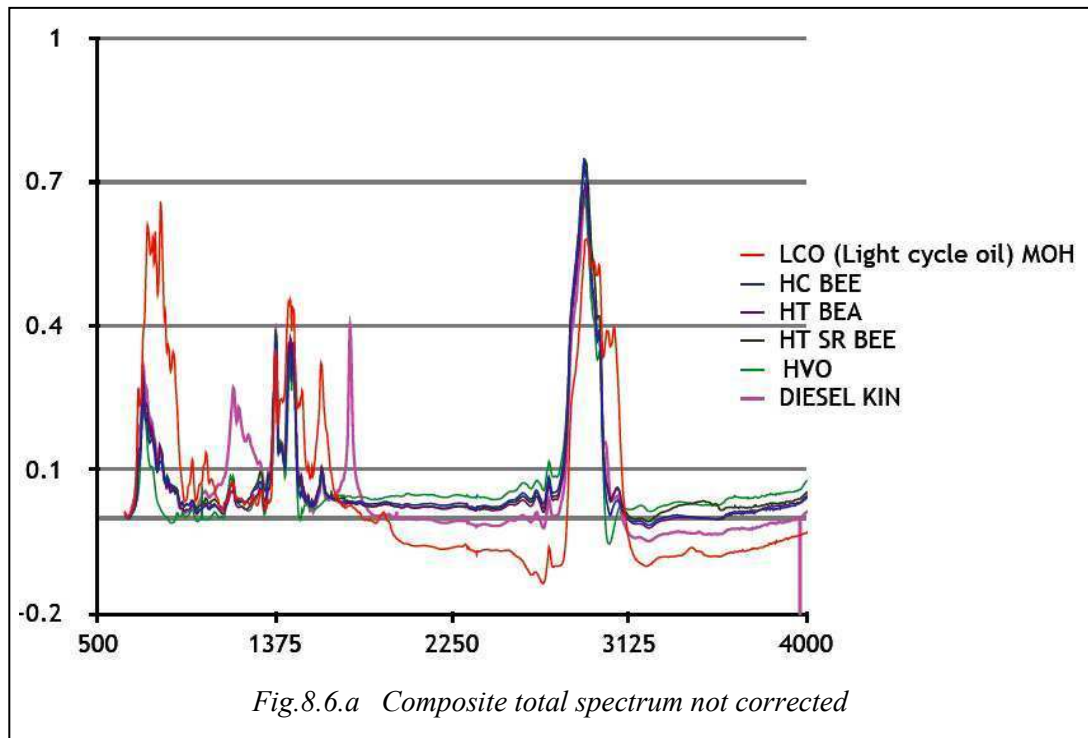




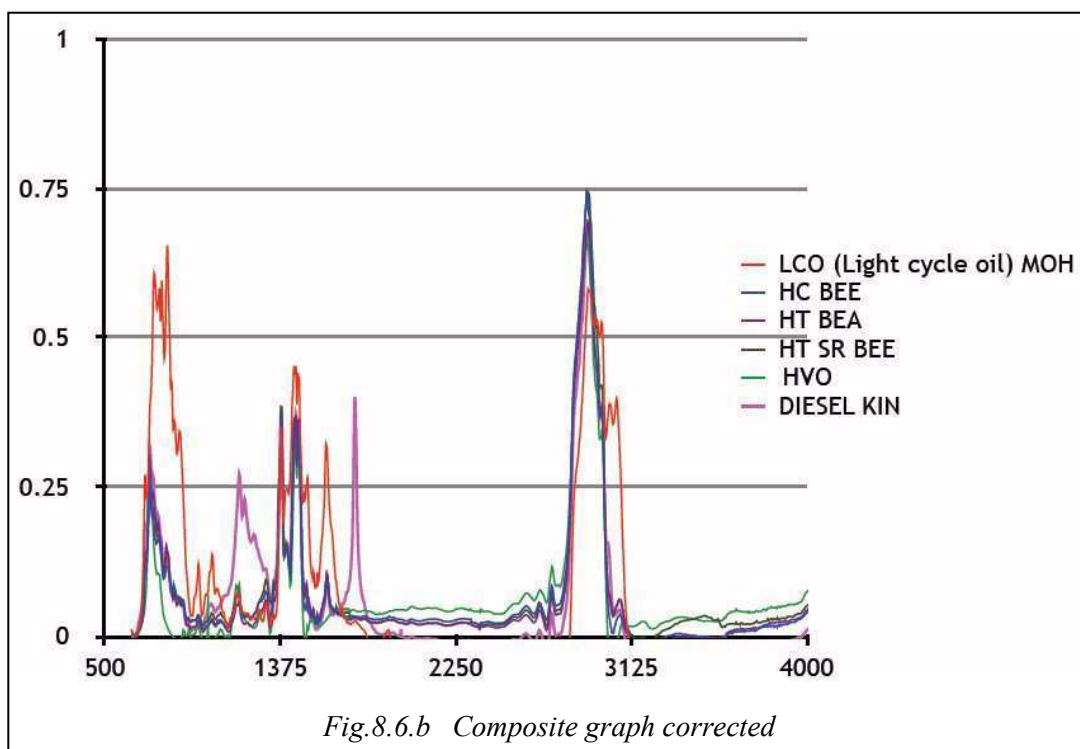
8.6 Comparison of diesel spectra

Obviously the peaks can not be analyzed by means of this analysis. They should be divided into equivalent regions. To facilitate the decision limits that will be introduced in the regions of the spectra the composite total spectrum will be used.

Following is the composite total spectrum of the five substances together.

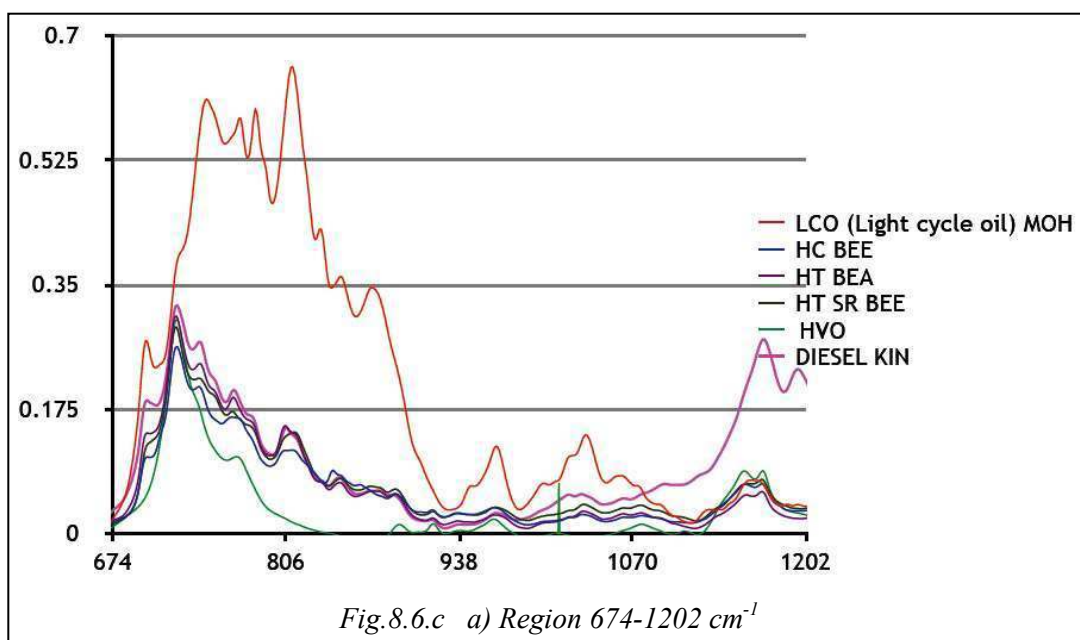


It is understood that errors occur in the HVO and to a much greater extent in the LCO. These errors are due to processor errors in converting the analog signal to digital. Besides, the functional region shown by the instrument is up to 1900 cm^{-1} . The identification can be done in this region. The total diagram corrected is illustrated as follows (*Fig.8.6.b*):



It is observed that the most important regions for analysis are:

- a) 674-1202 cm^{-1}
- b) 1202-1780 cm^{-1}
- c) 2609-3190 cm^{-1}



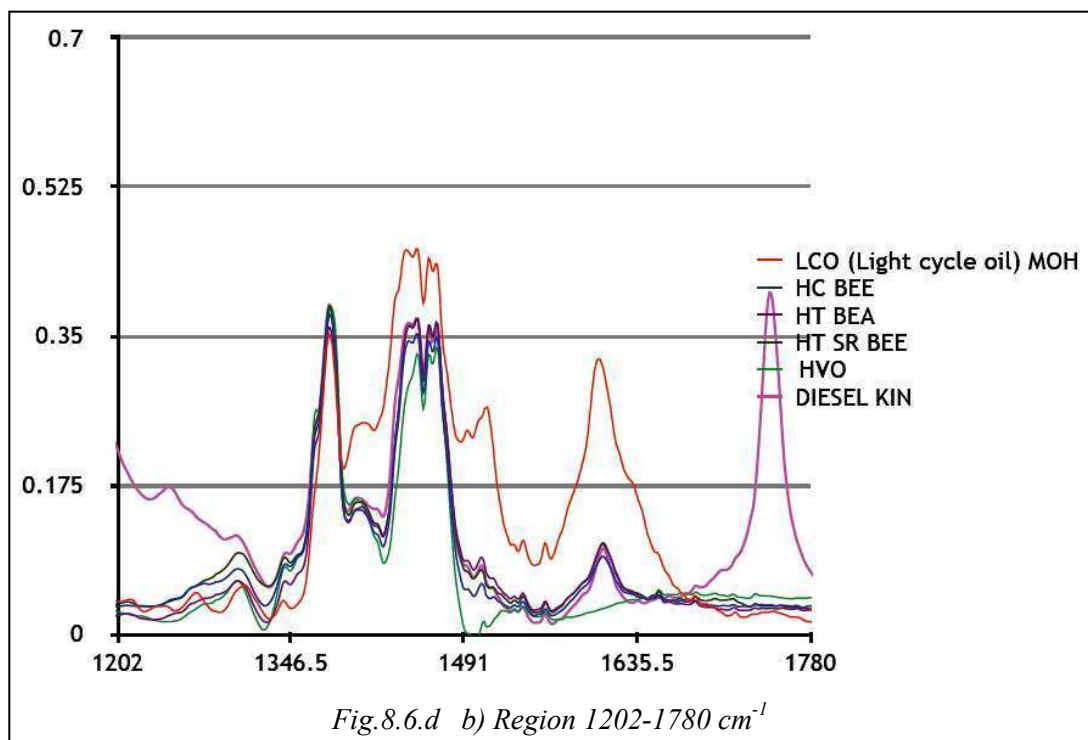
In this Figure (*Fig.8.6.c*) the wavenumbers or the ranges of interest are:

701 cm^{-1} , 724 cm^{-1} (all but LCO), 811-815 cm^{-1} , 970, 1037, 1158 and 1171 cm^{-1} .

First it is observed that the LCO shows huge differentiations compared with the other fuels, which seem to go hand in hand in most of the wavenumber range.

The corresponding substances that cause the absorptions are:

NO_2 bending, C-H out of plane or Double 1,3 aromatic or Triple 1,2,3 aromatic, Double 1,3/1,2 aromatic or Triple 1,2,3 aromatic, Aliphatic asymmetric P-O-C stretching bond (much higher concentration in LCO), Double S=O stretching bond, Asymmetric SO_2 stretching bond.



In this Figure (*Fig.8.6.d*) the wavenumbers of interest are: 1307, 1344, 1381, 1405, 1452, 1469, 1510 and 1608 cm^{-1} .

The corresponding interpretation is as follows:

Aliphatic hydrocarbon Methylene wagging, Aromatic C-N stretching. The point 1381 cm^{-1} is observed to be strong and same for all diesel types. It is characteristic Methyl symmetric C-H stretching. At 1405 cm^{-1} is a weak absorption =C-H in plane bending. At 1452 cm^{-1} is also an aliphatic hydrocarbon bond Methylene scissoring. 1469 cm^{-1} is pure Methyl asymmetric C-H stretching. The wavenumber 1510 cm^{-1} is due to stretching of the double bond C=C. Finally 1608 cm^{-1} is due to Pyridine C=N stretching and to C=C stretching bond.

What makes impression in this picture is the peak of motor diesel at about 1750 cm^{-1} . This

peak is characteristic since it is not found at all in the other spectra and is due to biodiesel contained in diesel.

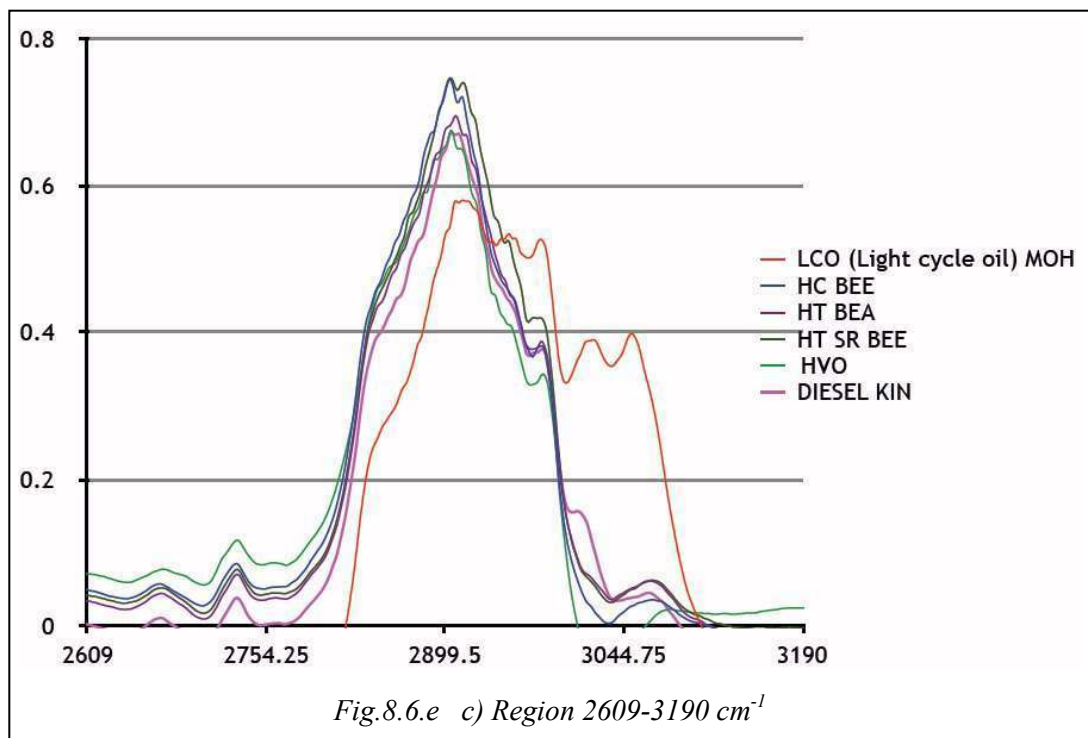


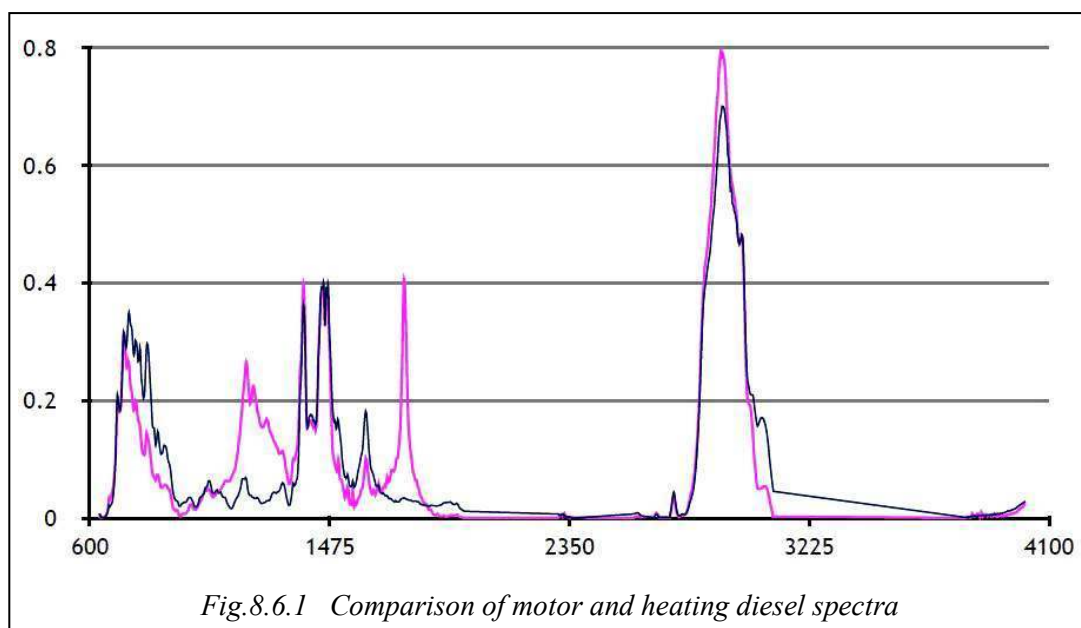
Fig.8.6.e c) Region 2609-3190 cm^{-1}

It is obvious that in this Figure (*Fig.8.6.e*) there are two peaks. One is located at 2905-2915 cm^{-1} and the other at 2980 cm^{-1} . In addition, the LCO has two peaks very intense. One is at 3022 and the other at 3054 cm^{-1} . The first is absent from the other types of diesel, while the other exists in very lower intensity.

The four peaks analyzed in sequence are as follows:

Carboxylic acid O-H stretching bond, Aldehyde C-H stretching bond, Carboxylic acid O-H stretching bond, Carboxylic acid O-H stretching bond.

8.6.1 Comparison of motor and heating diesel spectra



It is observed in the spectrum that the main differences are located at 1174 cm^{-1} and 1748 cm^{-1} .

As far as the absorption observed at 1174 cm^{-1} is concerned, it is due to Ether C-O stretching. Moreover it can be due to Aliphatic C-N stretching. It can also be due to SO_2 asymmetric stretching. The other peak observed universally in motor diesel is due solely to mixing with biodiesel.

9. Conclusions / Results

FTIR method offers many advantages for studies of this type. It is the most suitable for identifying chemical elements especially in organic compounds. This is because it takes advantage of fundamental movements of fundamental particles, i.e. even the slightest disturbance of the molecules give absorbance spectrum. This means that the method exhibits high sensitivity. This fact, combined with the rapid development of interferometers and processors give measurements remarkable accuracy. The interferometers depending on the length being traveled and construction materials reduce the analysis to only 1-2 cm⁻¹, while the processors convert the analog signal to digital more accurately.

The new statistical methods developed in recent years, in this case the cluster analysis and multi linear regression, have been enormously helpful in strengthening the infrared spectroscopy for identification of complex mixtures such as fuel. In modern times, the time is very important, and these methods contribute to its preservation since in just a few seconds are able to select the suitable spectrum from hundreds of others in a multivariate level.

The spectra obtained from fuels in the case of petrol and diesel as well, at about 3000 cm⁻¹ it was observed a very high peak with almost total absorption. In some cases, the absorbance exceeded the value of 1 which is considered an error. This way it was verified what was cited in the theory, that is, the noise-to-signal ratio plays a very significant role in the quality of the measurements. At very strong absorptions the signal reaching the detector is very weak and so the probability of error increases. Respectively, in ranges where no absorptions occur, the signal received by the detector is very strong resulting again errors in recordings.

Regarding the spectra in the experimental part, although they were taken from 631 to 4000 cm⁻¹, the ranges that were helpful in the examination and identification were at 600-1600 or 600-1800 cm⁻¹ or less. There, the spectrum was clear and continuous, did not go below zero or over the value of 1. That is why this specific range is called the "fingerprint" of the mid infrared. Here the spectrum is diversified as to all ingredients and hardly two spectra are identical in this region if they are not the same.

A difficulty in the comparison of spectra is the overlaps occurring in wavenumber values. That is, the wavenumber matches to many possible molecules. This is where the experience and observingness are needed, while at the same time helpful is what happened in the experimental part, i.e. "plotting" of spectra in a diagram.

The concept of calibration in this instrument should be that the calibration is continuous during all the years of its use. The more samples the user can input the more the measurement accuracy will be improved, because the properties of his standards will be related to the properties of the samples he will deal with. Therefore it would be beneficial whoever in the laboratory has samples of known properties in his possession, he should input them into the instrument's libraries.

An experimental result which is of great interest is the measurement of a gasoline sample by using the available libraries that were created during the experiments, and without them. Specifically, the New-MOH and HELPE libraries were removed and the results were as follows (*Fig.9.a* and *Fig.9.b*):

Measure Result System

29.09.14 14:54

SAMPLE ID 95R Prathriou
 DENSITY 0.7478 [g/ccm] @ 15.0 C
 COMMENT

	Parameter	Vol%	Mass%	
1	RON	96.2		MLR
2	MON	85.4		MLR
3	MTBE	5.04	5.00	
4	TAME	3.17	3.30	
5	Oxygen		1.42	
6	m-Xylene	5.08	5.89	
7	o-Xylene	3.90	4.61	
8	p-Xylene	2.82	3.26	
9	Toluene	10.0	11.6	
10	Olefins	11.3	10.9	MLR
11	Saturates	46.4	40.9	
12	Aromatics	34.2	39.9	MLR
13	Benzene	0.86	1.02	
14	2-Ethyltoluene	1.93	2.29	
15	Naphthalene	0.32	0.44	

Add Print Ok

Fig.9.a First results screen without the libraries

Measure Result System

29.09.14 14:54

SAMPLE ID 95R Prathriou
 DENSITY 0.7478 [g/ccm] @ 15.0 C
 COMMENT

	Parameter	Vol%	Mass%	
11	Saturates	46.4	40.9	
12	Aromatics	34.2	39.9	MLR
13	Benzene	0.86	1.02	
14	2-Ethyltoluene	1.93	2.29	
15	Naphthalene	0.32	0.44	
16	Oxygenates	8.21	8.30	
17	DCPD	0.0	0.0	
18	Di-Olefins	0.0	0.0	
19	AKI	90.8		
20	IBP	36.2		C MLR
21	T10	51.4		C MLR
22	T50	91.3		C MLR
23	T90	144.6		C MLR
24	FBP	192.1		C MLR
25	Maha. Distance	0.1		

Add Print Ok

Fig.9.b Second results screen without the libraries

Results with the libraries being available (Fig.9.c and Fig.9.d):

Measure Result System
29.09.14 14:51
SAMPLE ID 95R Prathriou
DENSITY 0.7479 [g/ccm] @ 15.0 C
COMMENT

	Parameter	Vol%	Mass%	
1	RON	95.3	CA4	
2	MON	85.0	CA4	
3	MTBE	5.08	5.04	
4	TAME	3.18	3.31	
5	Oxygen		1.43	
6	m-Xylene	5.09	5.91	
7	o-Xylene	3.91	4.62	
8	p-Xylene	2.78	3.21	
9	Toluene	10.2	11.8	
10	Olefins	10.3	10.0 CA4	
11	Saturates	50.1	45.1	
12	Aromatics	31.3	36.5 CA4	
13	Benzene	0.90	1.06	
14	2-Ethyltoluene	1.71	2.02	
15	Naphthalene	0.29	0.39	

Add Print Ok

Fig.9.c First results screen with the libraries

Measure Result System
29.09.14 14:51
SAMPLE ID 95R Prathriou
DENSITY 0.7479 [g/ccm] @ 15.0 C
COMMENT

	Parameter	Vol%	Mass%	
11	Saturates	50.1	45.1	
12	Aromatics	31.3	36.5 CA4	
13	Benzene	0.90	1.06	
14	2-Ethyltoluene	1.71	2.02	
15	Naphthalene	0.29	0.39	
16	Oxygenates	8.26	8.34	
17	DCPD	0.0	0.0	
18	Di-Olefins	0.0	0.0	
19	AKI	90.2		
20	IBP	31.3 C	CA4	
21	T10	47.0 C	CA4	
22	T50	86.2 C	CA4	
23	T90	148.7 C	CA4	
24	FBP	183.3 C	CA4	
25	Maha. Distance	0.0		

Add Print Ok

Fig.9.d Second results screen with the libraries

The displays are divided into first and second screens since the results do not fit on one screen, and the user has to move the cursor downward. As mentioned in the previous chapters where the instrument was described, the statistical method used for the calculation of a property is written next to the property in the results screen. Here it is observed that if the libraries are taken into account, CA4 is used which indicates very high accuracy in relation to the MLR used in other cases. Also there is a big difference in the calculation of octane number. Generally the convergence by using CA4 although theoretically is the worst of the CA, gives very good results. Of course it is very difficult to achieve CA1 unless the detector achieves the same standard by which the calibration took place.

Because in the field of gasoline more work was done in the sense that more samples were input, the results were improved significantly. In the field of diesel the differences are not yet appreciable, that is, the results are almost the same either the new libraries are used or not.

One difficulty that was encountered was the collection of samples. As known fuels never come out from the refineries without a cause, because of tax reasons. Therefore each time there was a small quantity of fuel taken as well as carefully to avoid any problem at the gateway. The laboratories staff in the refineries is constantly busy so the communication with them was limited which made the receipt of the samples even more difficult.

Perhaps the calibration that was made has not very good results, since 30 gasoline samples 95 R, 97 R and 100 R were input into the instrument. For example, only three 100 R samples were entered in the libraries, therefore making understandable that it is difficult for an unknown sample of 100 R gasoline to coincide with one of these 3. This is because the calibration was not with long term samples as for the refineries it is true that they switch suppliers or use different tanks. Finally it is understood that 3 samples are a very small number to represent the 100 R gasoline that will end up in the analyst's hands as a sample for investigation.

Samples entered in the libraries, were entered by the user two and three times each at least. This was done for two reasons: In order for the statistical methods of Eraspec to function, it is necessary for the same property e.g. the aromatics, that more than 40 samples are available so that the software made use of the same because there were not too many new samples. The second reason is for repeatability purposes. That is, if some small error occurred to any of the three samples, the backup measurements would reduce it significantly.

Samples must be entered quickly i.e. a few days after received from the refineries. Otherwise there is a risk of entering adulterated samples, i.e. matching of spectra to properties that do not correspond to reality. This can lead to a permanent status of displaying false results. Therefore it is necessary that the samples be entered to the libraries directly after they are opened, since e.g. the aromatics evaporate immediately.

10. Bibliography

1. Stuart, B. *Biological Applications of Infrared Spectroscopy*, ACOL Series, Wiley, Chichester: UK, 1997.
2. Wagner, C. On-site Fuel and Lubricant Quality Control by Portable FTIR Analysers. *Petro Industry News* [Online] **2013**, 34-35
http://www.petro-online.com/articles/analytical-instrumentation/11/dr._christoph_wagner/
3. Smith B. C. *Quantitative Spectroscopy: Theory and Practice*, Elsevier, Amsterdam (2002)
4. Mark, H. *Principles and Practice of Spectroscopic Calibration*, Wiley, New York, 1996
5. Hasegawa T. Principal Component Regression and Partial Least Squares Modeling. In *Handbook of Vibrational Spectroscopy*; Chalmers, J. M. and Griffiths, P. R., Ed.; Wiley: Chichester, UK, 2002; Vol. 3; p 2293–2312
6. Brown, S. D. Chemometrics. In *Encyclopedia of Analytical Chemistry*; Meyers, R. A., Ed.; Wiley, Chichester, UK, 2000, Vol. 11; p. 9671–9678.
7. Franke J. E., Inverse Least Squares and Classical Least Squares Methods for Quantitative Vibrational Spectroscopy. In *Handbook of Vibrational Spectroscopy*; Chalmers, J. M. and Griffiths, P. R. Ed.; Wiley, Chichester, UK, 2002; Vol. 3; p 2276–2292.
8. Yang H., Discriminant Analysis by Neural Networks. In *Handbook of Vibrational Spectroscopy*; Chalmers, J. M. and Griffiths, P. R. Ed.; Wiley, Chichester, UK, 2002; Vol. 3; p. 2094–2106.
9. Rusak D. A., Brown, L. M. and Martin, S. D., *J. Chem. Edu.*, **80**, pp. 541–543 (2003)
11. Peter R. Griffiths, James A. DeHasheth. In *Fourier Transform Infrared Spectrometry*; 2nd ed.; Wiley; New York; 2007.
12. Weisberg, S. *Applied Linear regression*, 3rd ed.; Wiley/Interscience (2005)
13. Consultation on the operation of the Eraspec from Dr. Philip Jordan Eralytics GmbH
14. Consultation on the operation of the Eraspec from Dr. Niklas Christenson Eralytics GmbH
15. Schartzmann, A Instruction Manual: In *Eraspec Fully Automated FTIR Fuel Analyzer*. Eralytics GmbH, Austria, Vienna 2013
16. Consultation on the operation of the Eraspec from Dr. Christoph Wagner Eralytics GmbH