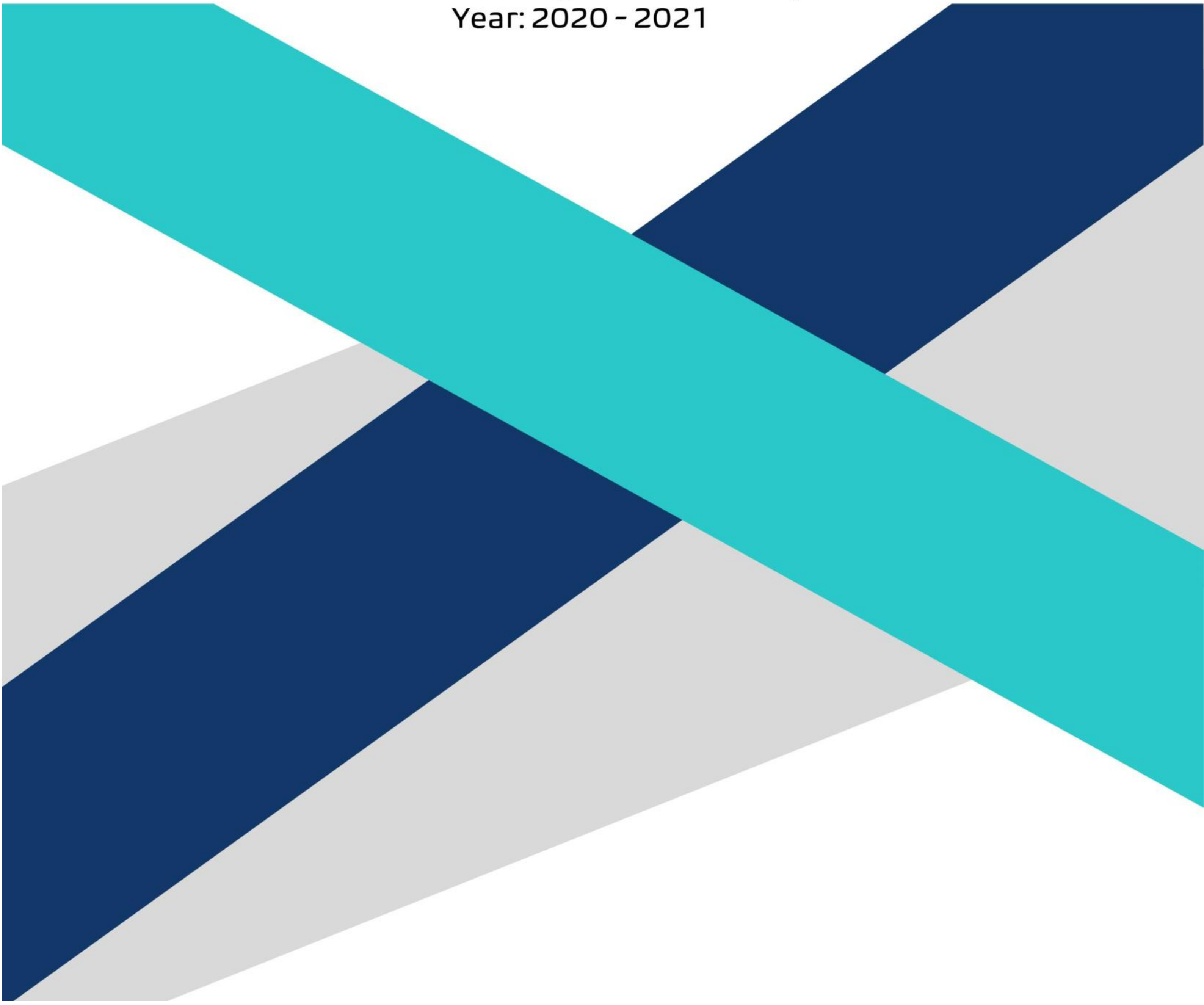
UNIVERSITAT ROVIRA i VIRGILI

QUÍMICA

BASF
The Chemical Company

Implementation of NIR Spectroscopy for the determination of physical and chemical properties in polyester samples by using Partial Least Squares (PLS) multivariate analysis

Student: Irene Molina Ruz
Academic Tutor: Òscar Pàmies
Professional Tutor: Marta Grau Carrasco
Bachelor's Degree in Chemistry
Year: 2020 - 2021



Index

1. Summary

The assessment of quality is one of the most important parts of the manufacturing process. To do so, a constant monitoring of the product properties has to be done. In the case of polyester resins, the acid value and the viscosity are the properties to control to assure the perfect production of the product. To achieve this goal, the implementation of NIR spectroscopy could be very useful. With this in mind, in this bachelor’s thesis, the objective is the application of NIR spectroscopy and the development of a Partial Least Squares (PLS) calibration method so that the determination of the properties mentioned above can be performed.

Garantir la qualitat d’un producte és una de les parts més importants d’un procés de producció. Per aconseguir-ho, s’ha de dur un seguiment constant de les propietats del producte. En el cas de les resines de polièster, l’índex d’acidesa i la viscositat son propietats que s’han de controlar per garantir una producció perfecta del producte. Per aconseguir monitoritzar aquestes propietats, la implementació de l’espectroscòpia NIR és molt útil. Amb aquesta idea, l’objectiu d’aquest treball de fi de grau, és l’aplicació de l’espectroscòpia NIR i el desenvolupament del mètode de calibratge Partial Least Squares (PLS) per realitzar la determinació de les propietats anteriors.

2. Objective

The aim of this bachelor's thesis has been the application of NIR spectroscopy to perform a follow-up of the polymeric products manufactured in the polyester resins plant in the complex of BASF Española S.L. in Tarragona. To assure the proper characteristics of the products, properties like acidity and viscosity must be constantly checked. For this reason, with the collected data, a multivariate model for each product has been developed in order to control these properties in an easy way without the need of tedious material.

3. Introduction

The development of this bachelor's thesis has been possible thanks to BASF Española S.L.[1] in their complex in Tarragona.

BASF is one of the biggest chemical companies in the world since it was founded in 1865 in Manheim, Germany. The company started to grow allowing the foundation of BASF Española S.L. in 1966. Three years later, the complex in Tarragona was inaugurated and since then, it has been expanding until becoming the biggest site in southern Europe.

BASF runs four different plants in the Tarragona complex, the biggest one is known as Cluster ED. This plant corresponds to the union of 3 different branches known as Dispersions, Polyesters, and Utilities. In Dispersions, the main production is based on acrylic dispersions used in the manufacturing of paints and adhesives. In the polyester plant, the main attention is focused on the manufacturing of polymeric resins that, as they present a high resistance to corrosion, they are used to produce pipes, pools and materials to be used in bathrooms and kitchens. However, lately, other compounds not defined as resins have been produced with the same reactants. The last branch of this plant, Utilities, oversees the treatment of all the waters used in all the industrial processes on the site.

In the case of the polyester plant, the synthetic procedure consists of the addition of alcohols (glycols) and acids into a reactor. The mixture is heated until it reaches the desired properties, the point at which the heating is stopped, and the mixture is cooled down. Afterwards, the product is solved in styrene.

As it was mentioned, the reactor is cooled down when the desired properties are obtained, for this reason, a constant monitoring of the mixture properties is required. At a certain time, a sample is taken from the reactor so that the acidity and the viscosity can be measured.

The extraction of samples from the reactor is done by an operator. This process may be dangerous due to the high temperatures of the reactors and the hazardous gases coming out of them. For this reason, the development of an in-line method that measures these properties without the need of extracting samples may be very useful. However, to finally obtain this goal, a huge library of samples must be collected, and a multivariate calibration model must be established for each product. The followed procedure and the results are explained in the coming pages.

4. Today's Chemical Industry

Nowadays, the chemical industry[2] is one of the largest and most diverse in the world. It grows exponentially each year as the manufacturing of new products is required for the new era humankind is facing. Many different sectors are found and some of them are just the first step of a large production process.

As a consequence of this growth, new industrial procedures have been developed. However, although the manufacturing procedures have evolved, they all rely on one main thing, the analytical control of the properties. In the last 50 years, analytical chemistry based on industrial processes[3] has grown in order to achieve a constant monitoring of the manufactured product with the objective of saving time, increase efficiency and provide more security. The information extracted can be also used to save energy and use less raw materials.

As the analytical chemistry has evolved, the properties measured have changed too. In the early years, these measurements involved only temperature and pressure data but, nowadays, information with a higher complexity is determined, like the composition or concentrations of one or more components. All of this has been possible to new analytical techniques and the adaptation of the existent ones to industrial processes. These innovations allow the automatization of analytical systems.

4.1. Batch processes

One of the most used chemical processes in the chemical industry is the one known as batch process. In this kind of industrial process, the reactants are introduced initially in the reactor that has a stirring system that favours the homogenization of the sample. Then, the temperature of the reactor is increased and, in some cases, catalysts and inhibitors are added. The main difference between this kind of process with others is that, once the product has been finished, the reactor is emptied completely so the process can start again with other reactants if necessary. This fact is the main advantage of this kind of reactors as they allow to carry process with different characteristics allowing, in this way, the obtention of a huge range of chemical products. The analytical control of these processes is trickier than others as they are evolving processes that cause a higher variability of the final product.

As a consequence, a constant monitoring of the process is required so that the properties at each time are known. This follow-up allows the verification that the process has been followed in the proper way.

4.2. Process control methodologies

The process monitoring can be done in different ways:

- **Off-line**: In this procedure, a sample is extracted from the reactor and is analysed in a laboratory. This kind of methodology allows the verification to be done by a professional with advanced material. However, the main disadvantage of this procedure is that a high amount of analysis is required and consequently much time is spent.
- **At-line**: In this methodology, a sample is taken from the reactor but in this case, the analysis is done near the production line. With this, the time spent is reduced but the main disadvantages are that the working atmosphere is not appropriate, the material used is less advanced and the person doing it is not a professional.
- **Online**: In this case, a sample is taken and transported to the automatised instrument. Once the sample has been conditioned, the analysis can be done in two different ways, in a discontinuous way (chromatographic methods) or in a continuous way (spectrophotometric methods).
- **In-line**: This procedure consists of the introduction of a sounding line in the reactors so that the measurements can be done without the need of extracting samples.
- **Non-invasive**: In this methodology, the contact between the analyser and the samples is eliminated completely.

The choice of the monitoring system depends on many factors of the production process. The need for the fast obtention of results, the available instruments, the cost of the analysis as well as the pre-treatment of the sample must be considered before making the choice.

5. NIR Spectroscopy

5.1. Introduction

In the year 1800, the astronomer Friedrich Wilhelm Herschel analysed the spectrum of sunlight. He directed sunlight through a glass prism so that the light was divided into its different colours. Herschel measured the heating ability of each colour using thermometers with blackened bulbs. When he measured the temperature just beyond the red part of the spectrum, he noticed some kind of invisible radiation. He found out that the area close to the red part had the highest heating ability of all. After this, he concluded that there had to be a different kind of light beyond the red portion of the spectrum that is not visible to the human eye. For this reason, Herschel decided to call this light "infrared (below red) light".

Herschel continued with his experiments, he placed a container filled with water between the prism and the thermometer. He observed that the temperature measured was lower than the one measured without the water container so he deduced that water must partially absorb the radiation. He also noticed that depending on the position of the prism, the temperatures measured for each colour varied too. This was the beginning of infrared spectroscopy which can be defined as the measurement of infrared light that a substance absorbs. This absorption depends on the wavelength of the light.

In the following years, many experiments were performed using Infrared Spectroscopy. In the late years of the 20th Century, the development of new electronics allowed the register of spectra in a fast and reproducible way. In Infrared Spectroscopy, in order to obtain data from the spectra, multivariate analysis techniques are needed. Nowadays, this technique is expanding as new chemometric techniques are being developed and the equipment is being improved.

5.2. Fundamentals

The Infrared Spectra region covers a wide range of wavelengths. Three different regions can be differentiated: FIR (Far InfraRed), MIR (Middle InfraRed) and NIR (Near InfraRed)[4]. The differentiation between regions is based on the interactions between light and matter as can be observed in Table 1.

Table 1. Regions of the electromagnetic spectrum.

Region	**Absorption Interval (nm)**	**Absorption Origin**
FIR	$4{\cdot}10^4$ - 10^6	Molecular Rotations
MIR	2500 – 40,000	Fundamental Molecular Vibrations

NIR	780 – 2,500	Overtones and Combination Bands of Molecular Vibrations

A molecule absorbs radiation when a change in the dipolar moment occurs as a consequence of the vibrational movement. Molecules are constantly moving, and, for this reason, two different vibrational movements can be differentiated, stretching vibrations and bending vibrations[5]. Stretching vibrations are found when there is a change in the distance between two atoms. Bending vibrations are found when a change in the angle between two atoms occurs. Among these two kinds of vibrations, another classification can be found as can be seen in Figure 1.

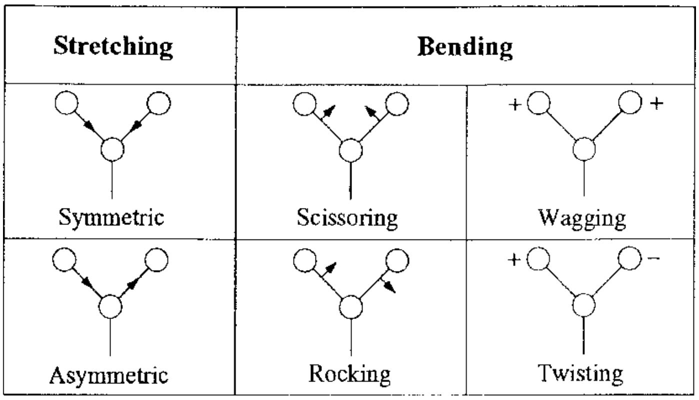


Figure 1. Stretching and bending vibrational modes.[5]

As the atoms in a molecule are never still, the movement between them is described as the simple harmonic oscillator[3,6]. In this approximation, a relation between the potential energy and the distance between two atoms is established. Potential energy is constantly changing approaching the minima when the two atoms are in equilibrium and being maximum when the distance between the atoms is maximum and minimum. This can be observed in Figure 2.

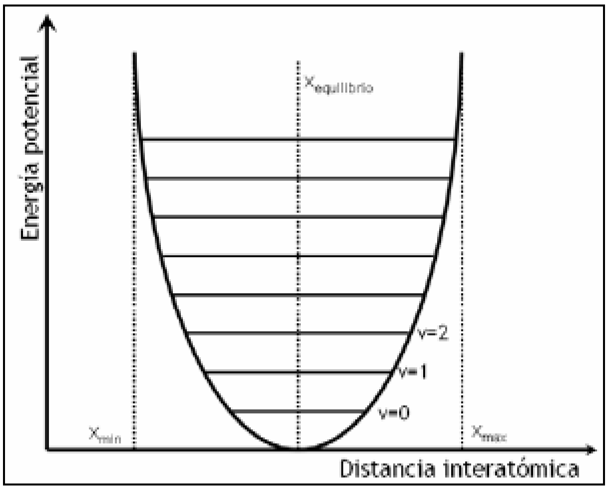


Figure 2: Harmonic oscillator potential energy

By applying the harmonic oscillator approximation[7], the potential energy can be defined as:

$$E = \frac{1}{2}kx^2 \qquad \text{Equation 1}$$

Where k is the force constant of the bond and x is the distance between atoms.

In a system formed by two masses joined by a spring, the vibrational frequency is defined as:

$$v = \sqrt{\frac{1}{2\pi}\frac{k(m_1+m_2)}{m_1m_2}} \qquad \text{Equation 2}$$

Where k is the force constant and $m1$and $m2$ are the masses of each of the atoms.

In classical mechanics, the behaviour of atomic particles is not described completely as it does not take into account the quantised nature of molecular vibrations. So, for an harmonic oscillator, the energy levels are calculated by means of Equation 3.

$$E_{vib} = \left(n + \frac{1}{2}\right)hv \qquad \text{Equation 3}$$

n is the quantic vibrational number (n = 0, 1, 2, …), h is the Plank constant and v corresponds to the vibrational frequency.

By combining Equation 2 and Equation 3, the vibrational energy of a diatomic molecule by taking into account its quantised nature can be defined as:

$$v = \left(n + \frac{1}{2}\right)\frac{h}{2\pi}\sqrt{\frac{k(m_1+m_2)}{m_1m_2}} \quad \text{Equation 4}$$

Despite this, the harmonic oscillator model is not appropriate to describe real molecules, because when two atoms approach, a repulsion between them is observed that increases the potential energy. In addition, when the distance between atoms increases, it approaches the distance at which the bond is broken and, as a consequence, the potential energy decreases. For these reasons, the behaviour of real molecules approaches an anharmonic oscillator. At low energetic levels, the energy curves for harmonic and anharmonic oscillators are quite similar and for this reason, it can be said that molecules show a harmonic behaviour when they are close to equilibrium. This can be observed in Figure 3.

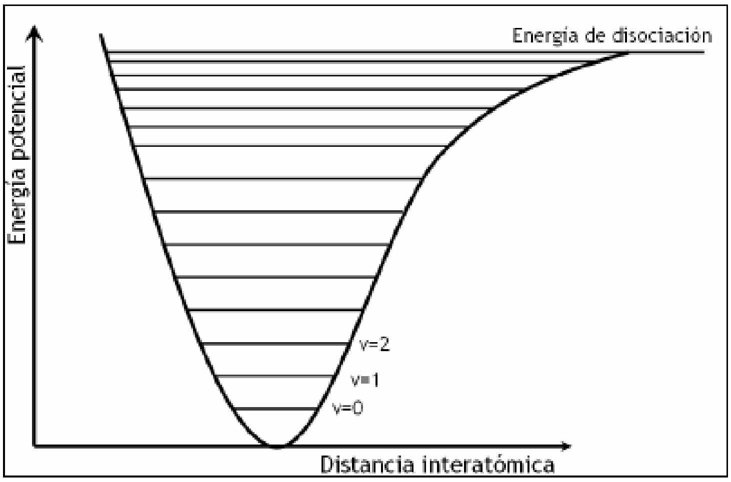


Figure 3: Anharmonic oscillator potential energy

The vibrational energy expression has to be corrected so that it explains the anharmonicity[7] of the molecules.

$$E_{vib} = \left(n + \frac{1}{2}\right) hv - \left(n + \frac{1}{2}\right)^2 hvy - \left(n + \frac{1}{2}\right)^3 hvyy' \ldots \text{Equation 5}$$

Where y and y' are the anharmonicity constants, n is the vibrational quantum number, h corresponds to the Planck constant and v is the vibrational frequency.

In an anharmonic system, the separation between levels is not the same and other bands different to the fundamental (Δn=$\pm$1) are found (Δn=$\pm$2, $\pm$3, …). These bands are known as overtones which are the ones that appear in the NIR region. The fundamental involving the absorption origin in NIR is based on the absorption of the NIR radiation when it is the same as the difference in energy between two vibrational levels and when there is a change in the dipolar moment of the molecule. The overtone bands correspond Δn>$\pm$1 although that depending on the type of bond, only the first and second overtones are visible. Combinations bands are found in polyatomic molecules as a consequence of energy changes of vibrational modes and these are also visible in the NIR spectra.

Overtones and combination bands have a lower probability to be seen than fundamental bands. As a result, NIR bands show a lower intensity than the ones in MIR. NIR bands tend to be wider and, in some cases, they cannot be identified properly due to overlapping of overtones and combination bands. In the NIR range, the intensity is related to the anharmonicity of the bond, the more anharmonic a bond is, the higher the probability of overtones and combination bands. For this reason, anharmonic bonds such as C-H, N-H and O-H, can be seen in NIR spectra. The common absorbent bonds in NIR can be observed in Figure 4.

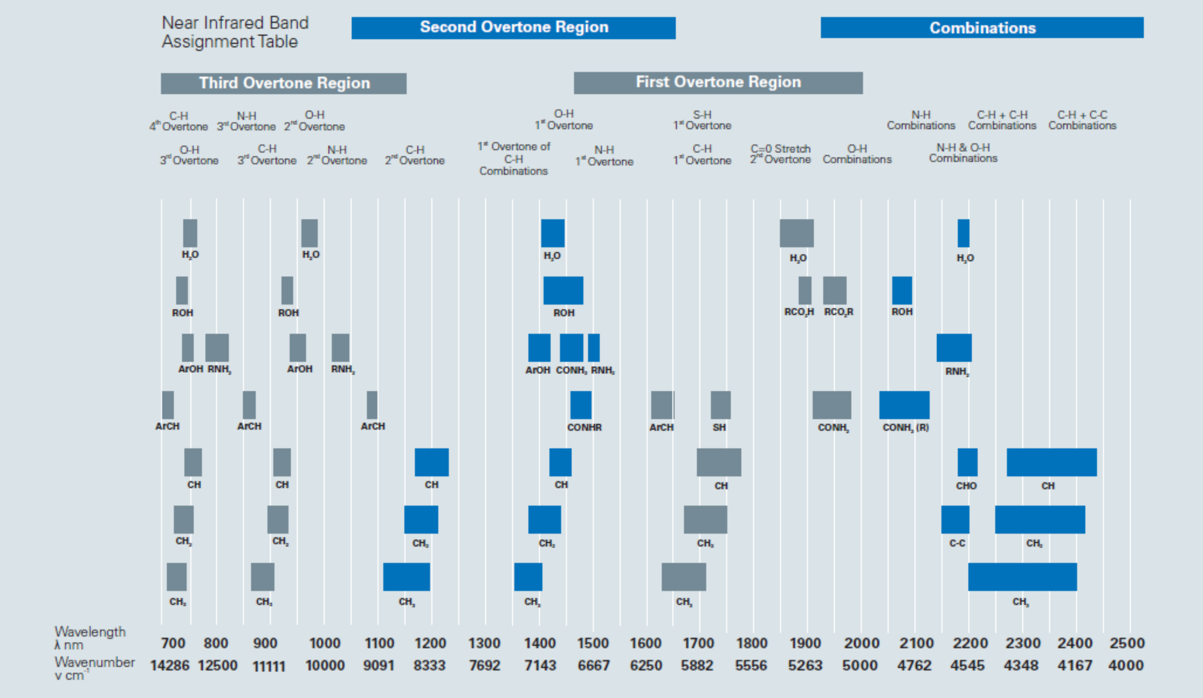


Figure 4: Common absorption bands in NIR.[8]

An important fact to take into account is that the interaction between molecules as well as hydrogen bonding, affect the NIR spectra by modifying the position of the bands. In addition, physical conditions like temperature, particle size as well as the crystalline structure of the sample affect the spectra too.

5.3. Instrumentation

The NIR equipment is quite similar to any other spectrometer[9]. However, as NIR bands are quite weak, proper instrumental stability and the elimination of noise must be an important thing to take into account. In the last years, the instrumentation of NIR spectroscopy has evolved as a response to the need to speed analyses. The devices used to record spectra are quite similar to the ones employed to measure electromagnetic spectrum, however, the NIR equipment requires devices to enhance analytical conditions[10].

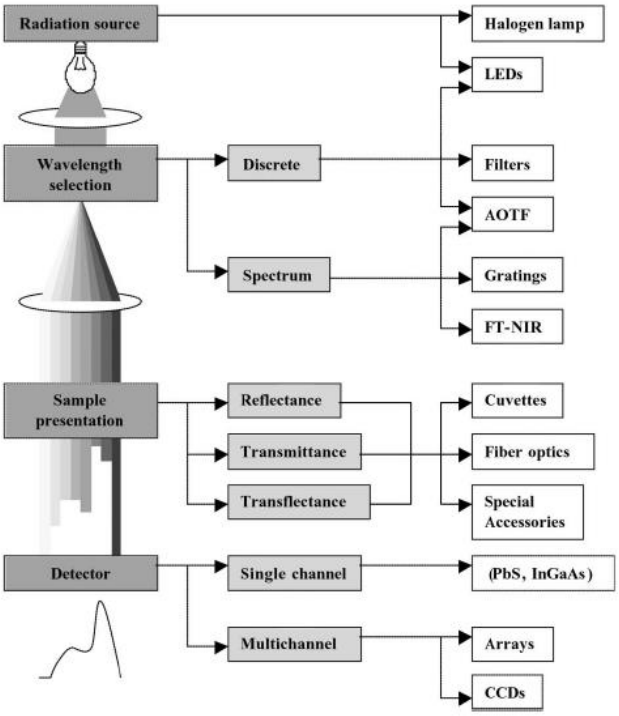


Figure 5: Principal features of NIR spectroscopy equipment.[10]

5.3.1. Radiation source

In NIR spectroscopy, the light source that is commonly used is the tungsten halogen lamp. Some of the characteristics that make this light source one of the most used worldwide are the robustness, the cost and the fact that provides a continuous spectrum from 320 to 2,500 nm. Other light sources that are used are LED (Light Emitting Diodes), which have a low working temperature and can emit until 1,600 nm.

5.3.2. Wavelength selection

Two types of NIR spectrophotometers can be differentiated according to the wavelength selection, discrete wavelength and whole spectrum.

In discrete wavelength spectrophotometers, wavelengths are selected by using filters as light sources which main goal is to allow the passage of variably broad wavelength bands or light-emitting diodes (LEDs) that emit narrow bands. These instruments are simple as they irradiate samples with only a few wavelengths. For this reason, they can only be used when working with analytes that absorb in specific spectral zones.

Whole-spectrum NIR spectrophotometers have a diffraction grating which makes them much more flexible than discrete-wavelength instruments, and, as a result, they can be used in many situations.

In recent years, new wavelength-selection devices have been incorporated into NIR spectrophotometers, *Acouto Optic Tunable Filters* (AOTFs) are an example. These use radio-frequency signals to alter the refractive index of a birefringent crystal to transmit at a single wavelength or to perform a wavelength scan faster than other designs. AOTFs are very suitable when used in production plants as they can stand aggressive conditions as they are not made by any moving parts.

5.3.3. Sample presentation

The way the sample is presented is an important part to take into account depending on the desired application. NIR spectroscopy is a technique that can be applied to different kinds of samples, solid, liquid and gaseous.

Three different measuring systems can be differentiated, reflectance, transmittance and transflectance. Reflectance is commonly used in solid samples, transmittance and transflectance are often used in liquid samples.

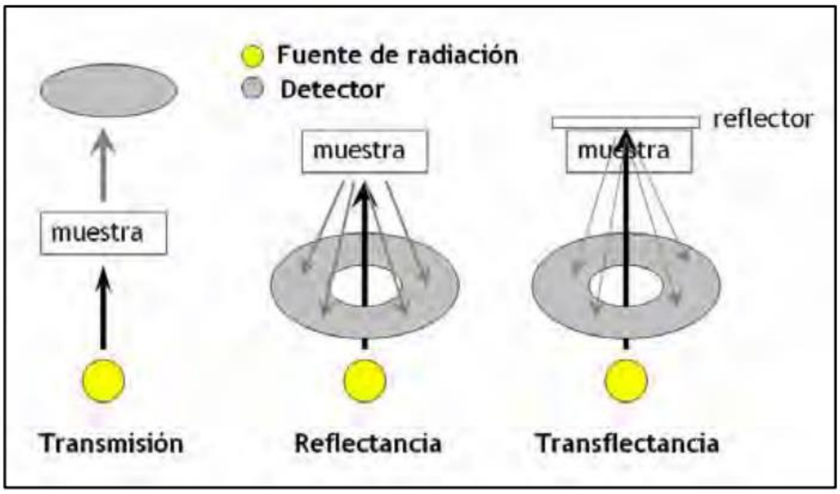


Figure 6: NIR measuring systems.[4]

5.3.4. Detector

In NIR spectroscopy, detection is performed by using devices made by semiconductors such as PbS or InGaAs. Two types can be differentiated, single-channel and multi-channel detectors. Multi-channel detectors consist of the arrangement of several detection elements placed in rows. In this way, several wavelengths can be measured at once increasing the speed of the analysis. These types of detectors have allowed the recording of spectra by using cameras that can determine composition at the same time it records the shape and size of the sample.

5.4. Advantages of NIR Spectroscopy

The implementation of NIR spectroscopy is a way to develop faster analytical methods avoiding, at the same time, the use of hazardous reagents. In addition, with NIR spectroscopy, a multicomponent analysis can be performed from one spectrum. For these reasons, chemical companies are trying to implement NIR spectroscopy on their chemical processes so that the quality of the final product can be assured. With this methodology, the production time could be reduced, analysts could be reduced, and less waste would be generated.

The in-line methodology can be very useful when optimizing experimental parameters such as temperature, the type of catalyst, etc. so that the process can be improved. Although the calibration methods are quite complex, once these are done, the analysis would become a routine that could be performed by inexperienced personnel.

6. Chemometrics

6.1. Introduction

In the last years, the appearance of modern and more sophisticated analytical techniques has led to the development of new types of equipment able to record bigger amounts of data. Then, this equipment is able to transform the data into useful information by means of mathematics and statistics. As a consequence, a new discipline known as Chemometrics[11] was created.

Chemometrics was defined in 1975 as an ensemble of methods used to extract useful chemical information from a group of original data. It uses mathematics and statistics to design optimum experimental procedures from analytical responses. Chemometrics[12] improve the efficiency of analytical procedures in all the stages and assures the quality of the final results. The term Chemometrics nowadays represents all multivariate calibration methods in analytical chemistry. Compared to the classical univariate calibration, this technique uses not only one spectral data point for the calibration, but the whole spectral structure.

As the NIR signal is quite complex, the use of chemometric techniques is needed to interpret and understand the data obtained from it. To obtain the desired results, quantitative or qualitative methods can be developed. To achieve this goal, the next steps were followed.

6.2. Stages on the development of multivariate calibration models

In a multivariate calibration model, experimental variables such as the NIR spectra are related to the known properties of the sample, in this case, acidity and viscosity. The objective of this calibration model is to predict these properties from unknown samples. To achieve this, a good calibration model has to be created by following the next steps:

6.2.1. Selection of a representative group of samples

The first step to achieve a good calibration model is to have a huge number of samples to develop the method. These samples must be representative of the variability during the process. They must include all the interval of the property to be determined and they must include the variability caused by other variables in the process such as the reaction temperature or the origin of the raw material.

6.2.2. Determination of the parameters to build the model

According to the American Society for Testing and Materials (ASTM)[13], the appropriate properties to monitor during esterification reactions are the acid value and the hydroxyl value. However, in this bachelor's thesis, the parameters used were the acid value and the viscosity. The analysis of these properties must be done with accuracy by means of appropriate analytical methodologies. This step is very important because the quality of the calibration model will depend on the analysed properties.

6.2.3. Obtention of the analytical signal

In the case of this bachelor's thesis, the measurements were performed by following an in-line methodology to obtain a huge library of spectra to be used to build the calibration model.

6.2.4. Data pre-treatment

When acquiring the data, undesired contributions could be observed in the spectra caused by the nature of the sample or the noise of the equipment. These can affect negatively the development of the calibration model. To avoid this, pre-treatment of the data is needed to reduce the spectral contributions of undesired effects and to improve the results of the model.

However, a statistical study of the results is performed by applying different pre-treatments. By doing this, the most appropriate pre-treatment is chosen. In conclusion, the pre-treatment is chosen after the study to avoid possible errors related to the sample or caused by unknown contributions.

Different mathematical data pre-treatment methods can be used when treating a group of samples[14]

6.2.4.1. Standard Normal Variate (SNV)

By using this pre-treatment, the multiplicative interference of scatter and particle size can be removed. Each spectrum is evaluated in an independent way for each spectrum and at each wavelength by following Equation 6.

$$SNV_i = \frac{r_i - \bar{r}}{\sqrt{\sum \frac{(r_i - \bar{r})^2}{(l-1)}}} \qquad \text{Equation 6}$$

SNV_i is the standard normal variate of a particular spectrum at a wavelength i, r_i the log 1/R at the same wavelength $\bar{r}$ is the average of all the log 1/R values at all the wavelengths in this particular spectrum, and l is the number of wavelengths.

6.2.4.2. Multiplicative Scatter Correction (MSC)

This pre-treatment[12] relates the scatter of each sample with respect to the scatter of a sample reference. As a consequence, the same level of scattering for all the spectra is obtained. To do it, the first thing is to select a reference spectrum which is usually the spectrum mean (r_m)

of all the spectra used for the calibration. Then each spectrum is corrected by applying Equation 7.

$$r = a + br_m + v \quad \text{Equation 7}$$

Where b corresponds to the slope indicating the influence of absorption multiplicative effects, a is the bias indicating a constant linear absorption additive effect with respect to the mean spectrum r_m and v is the residuals’ vector giving the differences between r and r_m. This residuals’ vector is assumed to contain the chemical variance in r. The new corrected spectrum $\hat{r}$ is obtained by means of Equations 8 and 9.

$$\frac{r-a}{b} = r_m + \frac{v}{b} \quad \text{Equation 8}$$

$$\hat{r} = r_m + \frac{v}{b} \quad \text{Equation 9}$$

6.2.4.3. First and Second Derivatives

By applying these pre-treatment methods, the peak overlapping is reduced, and the constant and linear baseline drift can be eliminated respectively. These are the ones most used in NIR spectroscopy to solve peak overlapping and baseline deviations. With this technique, details are more visible, and the result does not depend on the used spectral range. However, the noise is increased.

6.2.4.4. Straight Line Subtraction (SLS)

In a selected range of wavelengths[15], a calibration line is built and then it is subtracted from a specific spectrum.

6.2.4.5. Constant Offset Elimination

Constant Offset Elimination enhances the intrinsic matrix absorption of the material, thereby, enabling the extraction of significant information from the matrix of original spectra. This pre-treatment technique can only be used to horizontal baselines.

6.2.4.6. Centring

This kind of pre-treatment is usual to be applied before the development of the calibration model. The average of each variable is calculated and then it is subtracted to each point in order to obtain all the data-dependent on this average value.

$$X_{centered} = X_i - \bar{X} \quad \text{Equation 10}$$

6.2.4.7. Scaling and autoscaling

These pre-treatments are used to eliminate differences due to the scale of the variables. Equation 11 is used when the data is scaled, and Equation 12 is used when the data is autoscaled.

$$X_{scaled} = \frac{X_{original}}{S_{\ column}} \quad \text{Equation 11}$$

$$X_{scaled} = \frac{X_{original} - \bar{X}_{column}}{S_{\ column}} \quad \text{Equation 12}$$

Despite being different pre-treatment methods, some of them can be combined like in the case of the first derivate and the straight-line subtraction or the first derivative and the multiplicative scattering correction.

6.2.5. Development of the calibration model

The spectroscopic information obtained from the NIR spectra can be directly correlated to physical and chemical properties. As it was mentioned in previous sections, in this bachelor's thesis, the properties to be evaluated are the acid value and the viscosity of the samples. To do so, usually, two steps are required: the calibration of the method and the analysis to determine the value of an unknown sample. This can be observed in Figure 7.

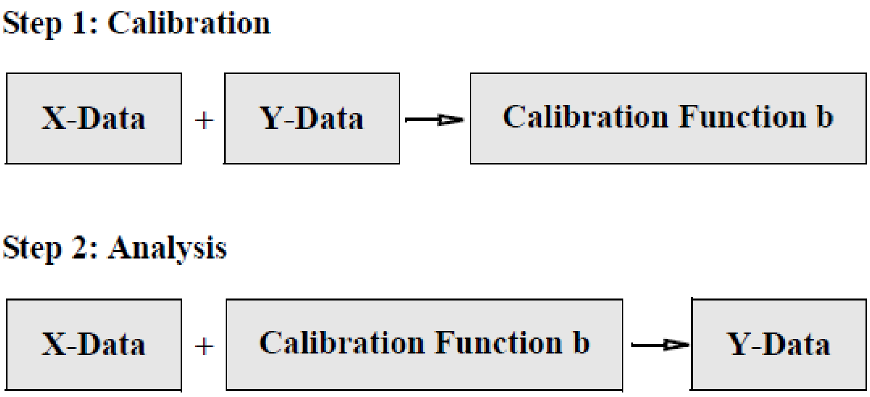


Figure 7: Schematic procedure of the quantitative determination of substances.[12]

Before building the calibration model, several parameters must be evaluated to avoid possible mistakes when computing. The frequencies intervals, abnormal samples, as well as the noise caused by the equipment must be taken into account.

Multivariate calibrations[16] make use of not only a single spectral point but take into account spectral features over a wide range. Therefore, the analysis of overlapping spectral bands or broad peaks becomes feasible. The information contained in the spectra of the calibration samples will be compared to the information of the acidity and viscosity values using a Partial

Least Square (PLS) regression. This method assumes that systematic variations observed in the spectra are a consequence of the change in the properties of the sample. However, the correlation between the properties and the change in the infrared signal does not have to be a linear one.

Multivariate calibrations require a large number of calibration samples and yield a large amount of data. In order to handle the data, the spectral data and the properties' data are written in the form of matrices, where each row in the spectral data matrix represents a sample spectrum and each row on the properties' data matrix contains the corresponding viscosity and acidity values of the samples. The matrices will be broken down into their Eigenvectors which are called factors or principal components. The advantage of this approach is that not all of the principal components are necessary to describe the relevant spectral features; for example, some of the vectors simply represent the spectral noise of the measurement. Only the relevant principal components will then be used instead of the original spectral data. The determination of the number of principal components is a crucial point for the quality of the calibration model. Using an insufficient number of principal components leads to poor reproduction of the spectral data and therefore the model will not be able to recognise changes in the spectral features, this is known as "underfitting". On the other hand, including too many principal components just adds spectral noise to the regression and does not increase the amount of valuable information, "overfitting".

In general, the aim of a quantitative analytical method is to determine the property Y of a system from an experimentally observable X, whereby X and Y are correlated by calibration functions b[16].

$$\vec{Y} = X \cdot \vec{b} \quad \text{Equation 13}$$

$$\begin{matrix} Y_1 & & Spectrum\ 1 \\ Y_2 & = & Spectrum\ 2 \cdot \vec{b} \text{ Equation 14} \\ Y_n & & Spectrum\ n \end{matrix}$$

The vector Y consists of the component values (of a single component) as determined by the reference measurements. The row vectors of the matrix X are formed from the calibration spectra. The aim is to determine the vector b, which is often called "regression coefficient".

$$\vec{b} = (X^T \cdot X)^{-1} \cdot X^T \cdot Y \quad \text{Equation 15}$$

When b is known, the prediction of unknown values for $Y_n$can be done.

Multicomponent systems can be analysed either for each component separately or simultaneously for all components. However, in this bachelor's thesis, the first option has been carried out as the analysis of each component separately usually yields better results.

6.2.5.1. Partial Least Squares (PLS)

6.2.5.1.1. Fundamentals

By using this calibration technique[16], the calibration of a single property can be done without taking into account other possible components that could cause variation. PLS is based on the fact the information can be contained in a smaller number of variables without any significant loss of information. For this reason, PLS is one of the most used calibration methods.

In order to perform a PLS regression, a comparison between the spectra and the analysed properties must be done. By doing this, changes in both data structures can be identified and correlated with each other. For this reason, a large number of samples must be analysed in order to obtain higher precision and error stability.

During PLS regression, the matrices X are reduced to only a few factors. In PLS, these factors are sorted in descending order, so the first factor gathers the main changes of the observed spectrum and for this reason, it is the most important factor for the calibration model.

During the calibration step, the PLS uses both the spectra data matrix (X) and the properties data matrix (Y). Both matrices can be decomposed as can be seen in Equations 16 and 17.

$$X = TP^t + F = t_1 p_1^t + t_2 p_2^t + t_3 p_3^t + \cdots + t_n p_n^t + F \text{ Equation 16}$$

$$Y = UQ^t + G = u_1 q_1^T + u_2 q_2^T + u_3 q_3^T + \cdots + u_n q_n^T + G \text{ Equation 17}$$

T and U are the scores matrices for each of the data, P and Q are the loadings for each of the data, and F and G, are the residuals of both data matrices.

The decomposition of both matrices is done in a simultaneous way, a relation is established between the scores of each matrix. By doing this, equation 18 has to be fulfilled.

$$u_n = b_n t_n \text{ Equation 18}$$

Where b_n is the regression coefficient between the scores of both matrices for each of the components. In this way, the new axis is used to obtain the maximum predictive ability for matrix Y.

Once the calibration model has been established, the prediction of new samples can be performed. To do so, the scores of the samples to be predicted (t*) are calculated by using Equation 19.

$$\hat{Y} = \sum_{a=1}^{A} b_n \, t_n^* q_n^t \text{ Equation 19}$$

By doing this, the value of the studied property is obtained.

6.2.5.1.2. Building of the PLS model and evaluation of the results

To build a proper PLS calibration model several points must be evaluated[3].

- **Selection of the number of factors**

  This part is a key point when using PLS because as it was mentioned in the previous section, the model has to contain the relevant information, but it has to exclude noise. The selection of factors is done by using a statistical parameter like the explained variance or PRESS.

$$PRESS = \sum_{i=1}^{N} (\hat{y}_i - y_i) \text{ Equation 20}$$

  In the case of this bachelor's thesis, the number of samples was quite small and for this reaction, the methodology used to validate the model was Cross-Validation. This method is explained in section 6.2.6.1. By using this, the Mean Square Error of Cross-Validation (MSECV) can be calculated.

$$MSECV = \frac{\sum_{i=1}^{M} (\hat{y}_i - y_i)^2}{M} \text{ Equation 21}$$

  Another way to establish the proper number of factors is to select the one that has a smaller predictive error. The Relative Standard Error (RSE) and the Root Mean Square Error (RMSE), are used to evaluate calibration and prediction errors defined as RSEC and RMSEC for calibration and RSEP and RMSEP for prediction.

$$RSE = 100\sqrt{\frac{\sum_{i=1}^{M} (\hat{y}_i - y_i)^2}{\sum_{i=1}^{M} y_i^2}} \text{ Equation 22}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{M} (\hat{y}_i - y_i)^2}{M}} \text{ Equation 23}$$

- **Explained variance percentage**

  This parameter determines how much information contained in the original variables is found in the new factors. Is one of the first parameters to be examined for both the spectral data and the properties data.
- **Bias**

  The bias is the average of the difference between the real values of the properties and the predicted ones. It can be calculated by using Equation 24.

  $$bias = \frac{\sum_{i=1}^{M}(y_i^{meas} - y_i^{pred})}{M} \quad \text{Equation 24}$$

  Where y_i^{meas} is the real value of the property, y_i^{pred} is the value predicted by the model and M is the number of analysed samples.
- **Loadings' plot**

  The loadings can be used to visually examine the factors that contain the noise of the calibration model.
- **Scores' plot**

  They are used to determine similarities and differences between samples, in this way, different clusters can be distinguished.
- **Mahalanobis distance**

  The Mahalanobis distance[17] can be defined as the difference between the measured spectrum of the sample and the values of all the spectra of the calibration set. The higher the difference, the higher the Mahalanobis distance. This can be caused by different factors like the contamination of the sample, or changes in temperature. The Mahalanobis distance can be calculated by applying Equation 25.

  $$h_i = t_i^T (X^T \cdot X^{-1}) \cdot t_i \quad \text{Equation 25}$$

  Where h_i is the Mahalanobis distance, t_i^T is the vector of matrix T, t_i is the vector of the original matrix, X^T is the T matrix and X is the original matrix.
- **Determination coefficient (R^2)**

  The determination coefficient is the proportion of the variance in the dependent variable that is predictable from the independent variable(s). The higher the value, the better the correlation between the values of the properties under study and the spectra.

$$R^2 = \left[1 - \frac{SEE}{\sum_{i=1}^{M}(Y_i - Y_m)^2}\right] \cdot 100 \text{ Equation 26}$$

Where SEE (Sum of Square Errors) is the sum of the residual values, Y_i is the value predicted by the model, Y_m is the average value of the predicted values.

- **Residual Prediction Deviation (RPD)**

  The RPD value[18] is the quotient of the standard deviation of the reference values (SD) and the standard error of Cross-Validation.

$$RPD = \frac{SD}{SECV} \text{ Equation 27}$$

The RPD value is a qualitative measure of the evaluation of the validation results. The higher the RPD, the better the calibration. In Table 2, a table with the optimum values of RPD can be observed.

Table 2: RPD optimum values classification[18].

RPD	Classification	Application
<1.0	Very poor	Not recommended
1.0 – 2.4	Poor	Not recommended
2.5 – 2.9	Right	Rough screening
3.0 – 3.9	Reasonable	Screening
4.0 – 5.9	Good	QC
6.0 – 7.9	Very good	QA
8.0 – 10.0	Excellent	Any application
>10.0	Superior	As reference

- **Residuals vs leverage**

  The leverage (h), can be defined as the spectrum x of the data matrix X in the following way:

$$h = x^t(X^tX)x \text{ Equation 28}$$

The leverage shows how different is the spectrum of a sample compared with the others. It is important to pay attention to samples with a high value of leverage because they could be outliers.

6.2.6. Validation of the calibration model

Once the model has been built, its predictive capacity has to be evaluated[12]. If undesired results are obtained, the model must be calibrated again. This evaluation is carried out by predicting a certain number of samples with a known value of acidity and viscosity with the chemometric model. The comparison of the predicted values with the real ones shows the precision of the model. Validation of different chemometric methods allows the recognition of outliers as well as the most suitable frequency ranges and allows the determination of the optimum number of factors.

There are two types of validation methods: internal validation (cross-validation) and external validation (test set validation).[12]

6.2.6.1. Cross-Validation

In this validation method, individual samples are taken from the calibration set. With the remaining samples, a chemometric model is established and used to analyse the samples taken out. By extracting these samples, they can be considered as independent and therefore they are not known to the calibration model. This process is done with the complete data set.

The comparison of the resulting analysis values with the original raw data allows the calculation of the predictive error of the complete data system, the Root Mean Square Error of Cross-Validation (RMSECV). The smaller the error, the better the quality of the model.

In the Cross-Validation, it's important to remove a small number of samples from the original data set to assure the remaining data set is similar to the original one. A scheme of the process can be observed in Figure 8.

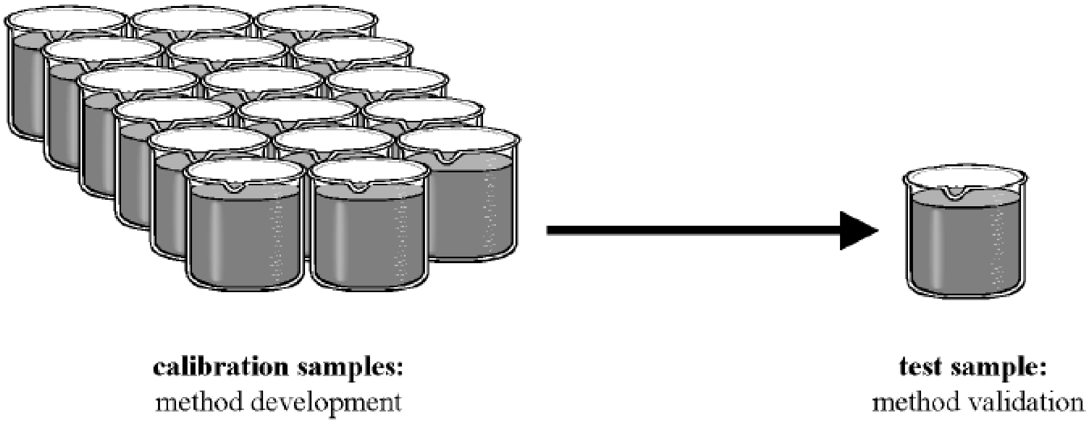


Figure 8: Schematic of a cross validation method.[12]

This methodology is mostly applied when there is a small number of samples.

6.2.6.2. Test set validation

In this validation method, in order to estimate the error of prediction for the model, more samples are measured to form a new group of samples called "test-set". The data set to perform this external calibration is divided into two groups, calibration samples and analysis samples (test-set). In this case, there is no exchange between the two sample sets. A comparison of the analysis results with the original data of the test set is used to calculate the Root Mean Square Error of Prediction (RMSEP). Figure 9 shows schematically how a test set validation works.

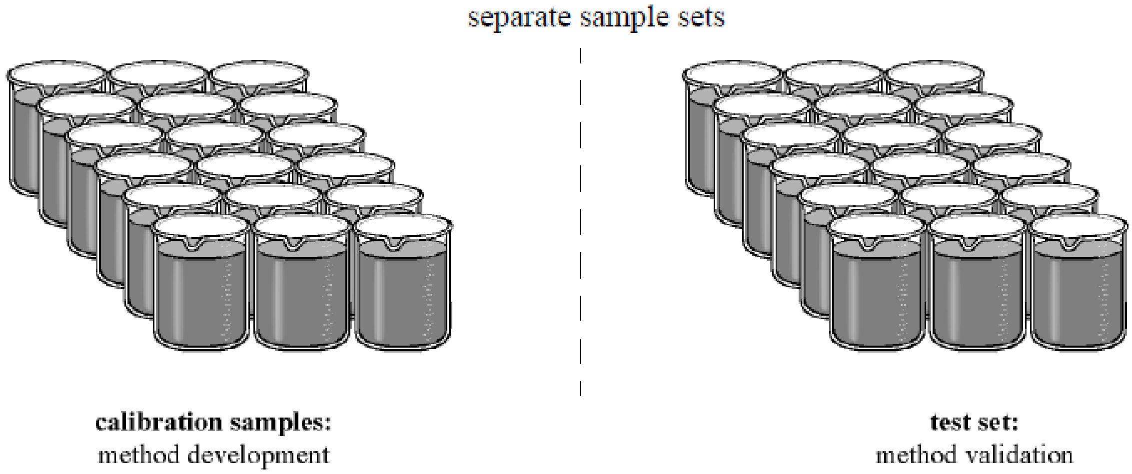


Figure 9: Schematic diagram of a test set validation.

This kind of methodology is mostly used when the number of analysed samples is big, in this way, two big groups can be differentiated.

7. Experimental Part

In this section, the methodology followed to perform the study on polyester resins is explained.

7.1. Polyester resins

The most important reaction in the production of unsaturated polyester resins is the esterification reaction. In this reaction, an acid and an alcohol react releasing water. This reaction is catalysed by acid and it is slightly endothermic.

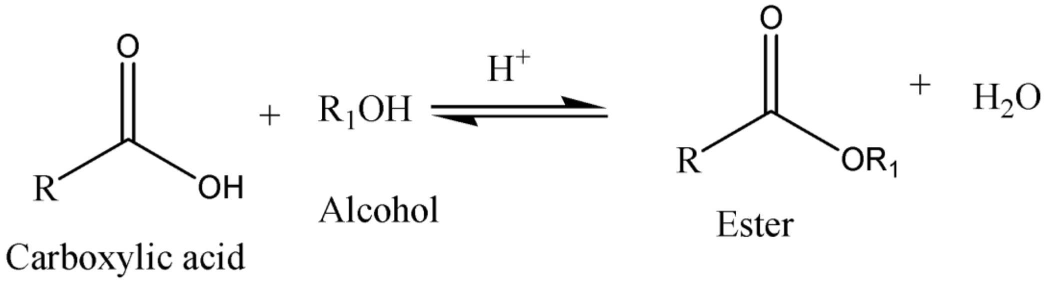


Figure 10: General esterification reaction.

In the production of unsaturated polyester resins, bifunctional acids (unsaturated and saturated) and bifunctional alcohols (glycols) are joined forming big chains. However, these chains are not linear but branched, as the glycols react with the double bond.

The production of these polyester resins can be divided into 5 steps.

1) The first step of the production process is the introduction of the glycols in the reactor. At the same time, an inert gas is fed into the reactor and the temperature is slightly increased.
2) Then, the acids are added to the glycols as well as the additives and the temperature is increased in a constant way.
3) The acids and the glycols react by means of an esterification reaction. In this reaction, water is released so it has to be eliminated from the reactor. The water that is being directly removed from the reactor contains a certain amount of unreacted raw material. The recovery of these unreacted raw materials is crucial to avoid the loss of reactants and to eliminate the possible hazardous gas emissions. To achieve this, the water goes through a distillation process.
4) Once the water has been removed, the condensation step begins. In this stage, the mixture is controlled by means of the extraction of a sample approximately every hour to measure the acid value and viscosity. To determine the endpoint of the reaction,

the viscosity is controlled. Once the viscosity value is appropriate to have a specific length of the chain, the reaction is stopped.

5) Then the mixture is transferred into another reaction filled with styrene. The mixture cannot be directly poured into the styrene because styrene would polymerise and for this reason, it has to be added bit by bit.
6) At this point, the production of the resin is done, however, in some cases, the resins can be modified with some additives so that can be used for a specific application. As a result, a wide variety of polyester resins can be obtained with different properties.

In the case of the polyester resin being studied in this bachelor's thesis, the reaction involves three types of glycols, monoethylene glycol, diethylene glycol and 1,2-polyethylene glycol that react with two acids, phthalic anhydride and maleic anhydride. Finally, the solvent used is styrene.

7.2. NIR Instrumentation

The data used to build and validate the calibration model was obtained by means of an in-line methodology. A sounding line was introduced in the reactor at a medium position so that it was in constant contact with the manufactured product. Approximately, every hour a sample was extracted from the reactor so that the acidity and the viscosity could be measured. By doing this, control of the polymerization reaction can be done.

To perform an in-line follow-up, a Matrix Fibre[19] (Matrix-F) which is a Fourier Transform Near-Infrared (FT-NIR) spectrometer was used. To this equipment, several sounding lines can be attached so that the monitoring of several reactions can be performed. The analysis is performed with 16 scans in a 4,000 to 12,000 cm^{-1} interval.

Figure 11: Matrix-F spectrometer with a sounding line.

7.3. Analytical measurement of samples

As It was mentioned in previous sections, the acid value and the viscosity were measured by following these analytical measurements.

7.3.1. Acid value measurement

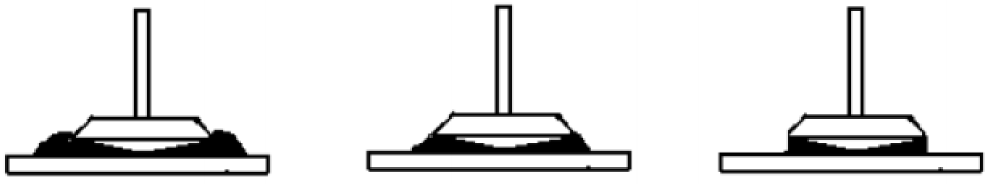
The fundamental of this technique consists in the neutralization of the free acids in the sample by means of titration with a 0.1 M potassium hydroxide solution. Then, the determination of the endpoint of the titration is done potentiometrically. The result is expressed in milligrams of potassium hydroxide needed to neutralise one gram of sample. To do this measurement, a specific amount of sample is weighted according to the expected value of acidity. This expected value is of acidity can be found in the literature of the compound. Then, this amount of sample is solved in a methanol: water solution 100:5 and tetrahydrofuran. The equipment shows the acid value by applying Equation 29.

$$Acid\ value = \frac{mL\ KOH\ used \cdot 5.61}{weighted\ amount\ of\ sample} \text{ Equation 29}$$

7.3.2. Viscosity measurement

The method used to determine the viscosity of polyester samples during the condensation is the one that uses a Cono-Plato Model C&P Brookfield CAP viscosimeter. Two parts can be differentiated in the viscosimeter, the cone and a sheet. This last part has the role of heating the sample so that it is maintained at a constant temperature.

The procedure goes as follows, first, the base is adjusted to the appropriate temperature for ten minutes, at this point, the cone and the sheet must be touching each other. This step is very important because polyester resins solidify quite quickly so it's crucial to keep the sample at a stable temperature so that it remains fluid. Then, the cone is lifted, and a small amount of sample is dropped in the sheet and after that, the cone is lowered again. The appropriate amount of sample can be seen in Figure 12.



Too much Appropriate amount Not enough

Figure 12: Appropriate amount of sample to assure the proper measurement of the Cono-Plato viscosimeter.

The dynamic viscosity of the sample is assessed by using the force that a liquid opposes when subdued to an effort done by a rotational body, the cone. The force is calculated by the apparatus and it is expressed in mPa·s.

7.4. Development of the method

The measurement of the spectra, the treatment of the data and the construction of the calibration model were done by using the software and methodologies explained in the next sections.

7.4.1. Software used to treat the data

7.4.1.1. CMET

This software[20] allows the control of the NIR spectrometer. It allows the development of analysis routines. It has been used to monitor the spectra. The software allows the creation of scenarios to automatise the analysis of samples. The scenario used in this bachelor’s thesis was built by the informatics team due to the complexity to link the scenario to the control screen in the plant.

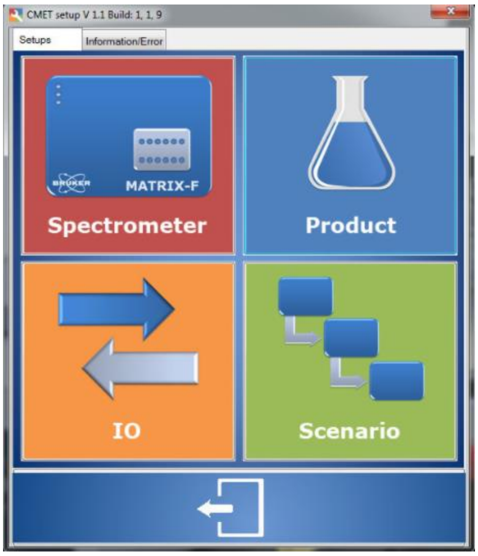


Figure 13: CMET set-up screen.

7.4.1.2. OPUS

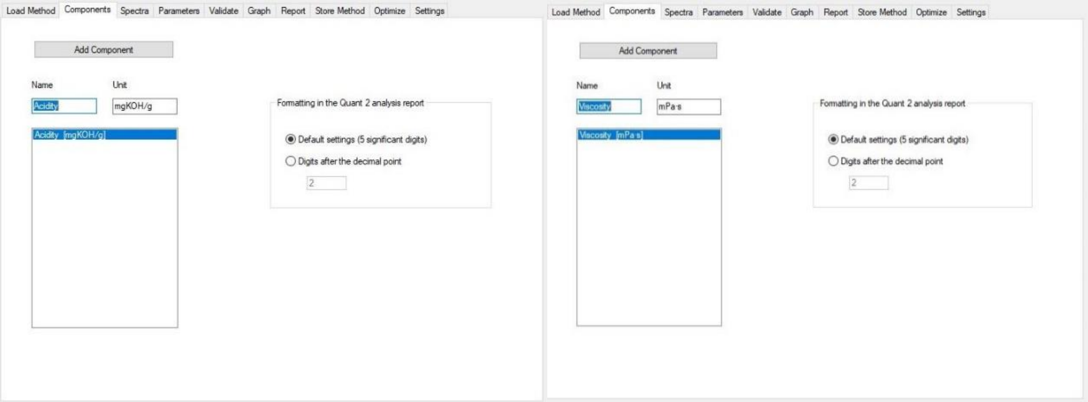
The OPUS software[21] distributed by Bruker allows the visualization of the recorded spectra and their treatment. In this software, the calibration of the method is performed by using the Quant 2 option.

7.4.2. Procedure used to develop the method

By means of the OPUS software, the calibration of the method was possible. Every step followed to achieve the desired results is explained in the following points.

1) Establishment of components of interest

In the components tab, the property to be analysed is introduced with the appropriate units. In this bachelor’s thesis, two separate methods were built each of them with one of the properties to be studied.



Name: Acidity
Unit: mgKOH/g
Acidity [mgKOH/g]

Formatting in the Quant 2 analysis report
● Default settings (5 significant digits)
○ Digits after the decimal point
2

Name: Viscosity
Unit: mPa·s
Viscosity [mPa·s]

Formatting in the Quant 2 analysis report
● Default settings (5 significant digits)
○ Digits after the decimal point
2

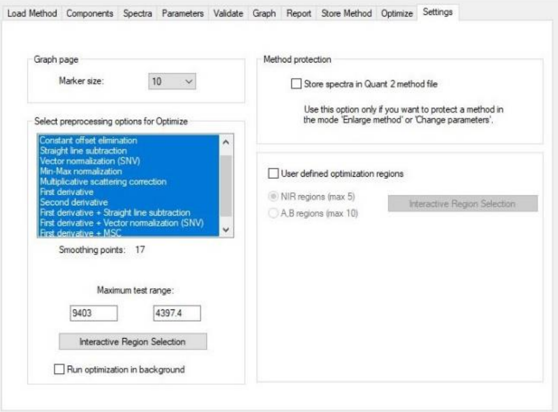
Figures 14 and 15: Components’ tab of the calibration method software.

2) Uploading of the spectra

The spectra of the product are introduced in the spectra tab. To each spectrum, the analysed value of the properties is introduced.

3) Optimization parameters

In the settings tab, the parameters to perform the optimization are selected. The pre-treatments to be used in the method and the region of the spectra to be analysed are chosen.



Graph page
Marker size: 10

Select preprocessing options for Optimize
Constant offset elimination
Straight line subtraction
Vector normalization (SNV)
Min-Max normalization
Multiplicative scattering correction
First derivative
Second derivative
First derivative + Straight line subtraction
First derivative + Vector normalization (SNV)
First derivative + MSC
Smoothing points: 17

Maximum test range: 9403 4397.4

☐ Run optimization in background

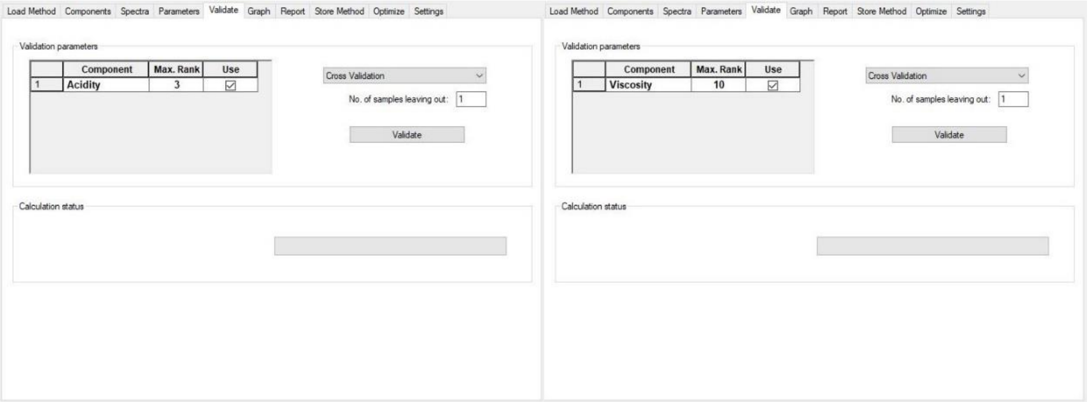
Method protection
☐ Store spectra in Quant 2 method file
Use this option only if you want to protect a method in the mode 'Enlarge method' or 'Change parameters'.

☐ User defined optimization regions
● NIR regions (max 5)
○ A,B regions (max 10)

Figure 16: Settings tab of the calibration method software.

4) Validation

The next step is to validate the method in the validation tab. As it was mentioned in previous sections, the validation method used was cross-validation. Then, the software performs the validation and the results are displayed in the graph tab.



	Component	Max. Rank	Use
1	Acidity	3	☑

Cross Validation

No. of samples leaving out: 1

	Component	Max. Rank	Use
1	Viscosity	10	☑

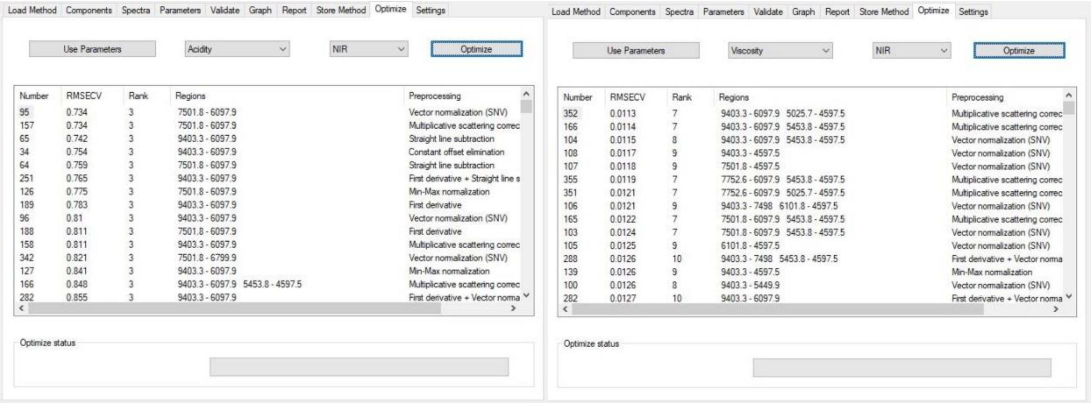
Cross Validation

No. of samples leaving out: 1

Figure 17 and 18: Validation tab of the calibration method software.

5) Optimization

The software performs several optimizations so that the best one can be used in the method. The best one is selected by looking at the RMSECV and RMSEP, the lower the value, the better the optimization.



Acidity

NIR

Number	RMSECV	Rank	Regions	Preprocessing
95	0.734	3	7501.8 - 6097.9	Vector normalization (SNV)
157	0.734	3	7501.8 - 6097.9	Multiplicative scattering correc
65	0.742	3	9403.3 - 6097.9	Straight line subtraction
34	0.754	3	9403.3 - 6097.9	Constant offset elimination
64	0.759	3	7501.8 - 6097.9	Straight line subtraction
251	0.765	3	9403.3 - 6097.9	First derivative + Straight line s
126	0.775	3	7501.8 - 6097.9	Min-Max normalization
189	0.783	3	9403.3 - 6097.9	First derivative
96	0.81	3	9403.3 - 6097.9	Vector normalization (SNV)
188	0.811	3	7501.8 - 6097.9	First derivative
158	0.811	3	9403.3 - 6097.9	Multiplicative scattering correc
342	0.821	3	7501.8 - 6799.9	Vector normalization (SNV)
127	0.841	3	9403.3 - 6097.9	Min-Max normalization
166	0.848	3	9403.3 - 6097.9 5453.8 - 4597.5	Multiplicative scattering correc
282	0.855	3	9403.3 - 6097.9	First derivative + Vector norma

Viscosity

NIR

Number	RMSECV	Rank	Regions	Preprocessing
352	0.0113	7	9403.3 - 6097.9 5025.7 - 4597.5	Multiplicative scattering correc
166	0.0114	7	9403.3 - 6097.9 5453.8 - 4597.5	Multiplicative scattering correc
104	0.0115	8	9403.3 - 6097.9 5453.8 - 4597.5	Vector normalization (SNV)
108	0.0117	9	9403.3 - 4597.5	Vector normalization (SNV)
107	0.0118	9	7501.8 - 4597.5	Vector normalization (SNV)
355	0.0119	7	7752.6 - 6097.9 5453.8 - 4597.5	Multiplicative scattering correc
351	0.0121	7	7752.6 - 6097.9 5025.7 - 4597.5	Multiplicative scattering correc
106	0.0121	9	9403.3 - 7498 6101.8 - 4597.5	Vector normalization (SNV)
165	0.0122	7	7501.8 - 6097.9 5453.8 - 4597.5	Multiplicative scattering correc
103	0.0124	7	7501.8 - 6097.9 5453.8 - 4597.5	Vector normalization (SNV)
105	0.0125	9	6101.8 - 4597.5	Vector normalization (SNV)
288	0.0126	10	9403.3 - 7498 5453.8 - 4597.5	First derivative + Vector norma
139	0.0126	9	9403.3 - 4597.5	Min-Max normalization
100	0.0126	8	9403.3 - 5449.9	Vector normalization (SNV)
282	0.0127	10	9403.3 - 6097.9	First derivative + Vector norma

Figure 19 and 20: Optimization tab of the calibration method software.

8. Results and Discussion

After following steps explained in the preceding parts, the calibration method was built, and the results are explained in the coming parts.

8.1. Study of the spectra

The visualization and analysis of the spectra are the most important parts to develop a consistent calibration method. In the NIR spectra, the overtones and combination bands of the bonds in the molecule can be observed. By using the band assignment tables[8], each peak can be associated and consequently, a variation in the properties of the product can be studied. In Figure 21, the spectra of the studied product can be observed.

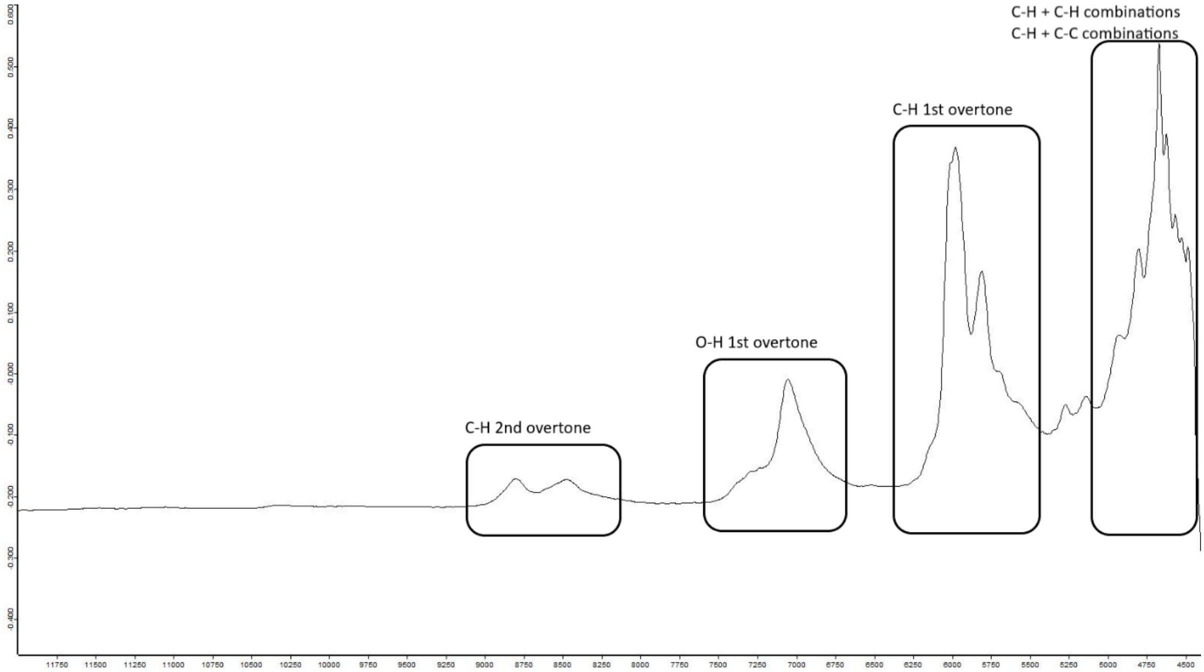


Figure 21: Spectra of the studied product.

In this spectrum, several peaks can be differentiated as is shown in the picture. In the 9,000 to 8,250 cm^{-1} spectral range, two small peaks can be observed corresponding to the second overtone of the carbon-hydrogen bond. In the range going from 7,250 to 6,750 cm^{-1}, a peak is observed which corresponds to the first overtone of the oxygen-hydrogen bond. The first overtone of the carbon-hydrogen bond can be observed in the spectral range going from 6,250 to 5,500 cm^{-1}. In the right part of the spectrum, several peaks are observed in a range going from 4,750 to 4,000 cm^{-1}, which might correspond to the combination bands of the carbon-hydrogen bonds and the combination bands of the carbon-hydrogen bonds with the carbon-carbon bonds.

8.2. Calibration methods

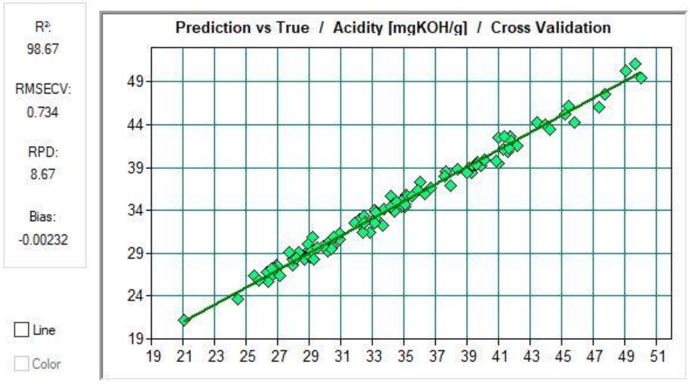
In order to build a good calibration method, the steps explained in previous sections were followed. In the end, two chemometric models were built, one for the acid value and another one for the viscosity.

8.2.1. Acid value calibration method

To develop this calibration method, 114 samples were measured and 101 were finally used to build the calibration method. Each sample was introduced in duplicate so that in the end, a library of 228 spectra could be obtained.

The first step to build the calibration model was to select the spectral region to be used and the optimization parameters. In this method, all the optimization parameters were selected, and the used spectral region went from 9,403 to 4,397.4 cm^{-1} as the region above 9,000 cm^{-1} is a spectral noise region and it can affect the calibration. This can be seen in Figure 16.

As the number of samples was not big enough to perform a test set validation, a cross-validation was performed to see which points were recognized by the method as outliers by using the Mahalanobis distance. In Figure 22, the calibration line of the method can be observed.


R²: 98.67

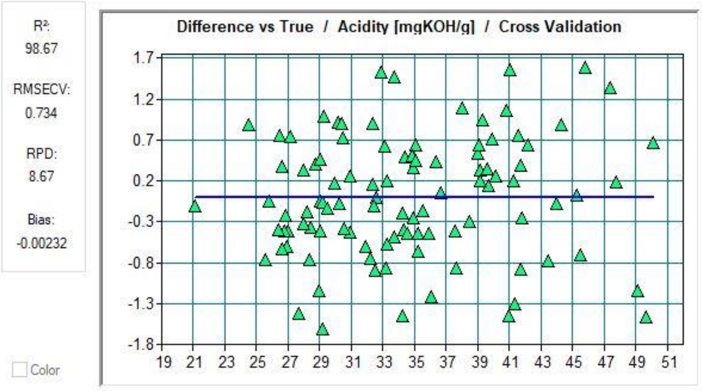
RMSECV: 0.734

RPD: 8.67

Bias: -0.00232

Figure 22. Calibration line of the acid value calibration method.

By taking a look at the difference between the predicted value and the real value plot (Figure 23), a blue line can be seen. This blue line is the average of the difference between the real values and the predicted ones. It can be seen that the points are distributed in a homogeneous way in the upper and lower part of the line, this means that there is no tendency nor correction factor in the prediction of the developed model.


R²: 98.67

RMSECV: 0.734

RPD: 8.67

Bias: -0.00232

Figure 23: Predicted value vs real value plot of the acid value calibration model.

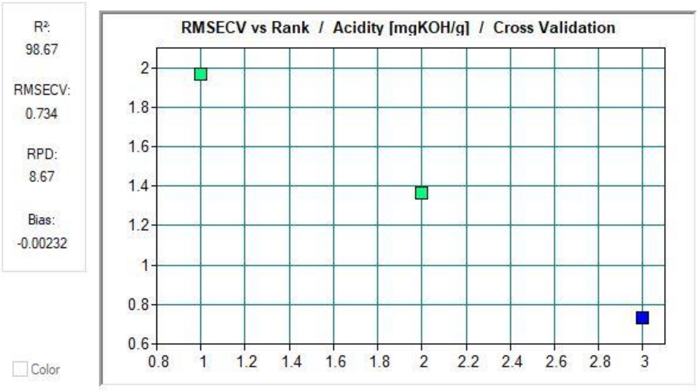
Table 3 gathers a summary of the characteristics of the acid value calibration model.

Table 3: Results of the cross-validation in the acid value calibration model.

R^2	RMSECV	RPD	Bias	Rank	MD	Region	Pre-treatment	Calibration Samples	Total Samples
98,67	0,734	8,67	-0,00232	3	0,08	7501,8 - 6097,9	SNV	101	114

It can be seen by looking at the Mahalanobis Distance value that the calibrated method is quite restrictive. As a consequence, the software recognised several outliers. This might seem like a big number, however, in some cases, problems with the reactor happened and for this reason, the spectra were not representative, so they had to be eliminated. Despite this, in some cases, the elimination of outliers does not lead to a better calibration model. In some instances, the software considers certain points as outliers because no other points are found in the same region.

The rank chosen by the software after the optimization is good as it gathers the necessary information contained in the factors. This can be observed in Figure 24.


R²: 98.67

RMSECV: 0.734

RPD: 8.67

Bias: -0.00232

Figure 24: RMSECV vs Rank of the acid value calibration model.

When optimizing the method, the spectra region the software has selected in the cross-validation corresponds to the region where the 1st overtone of the oxygen-hydrogen bond and the pre-treatment to be used is the Standard Normal Variate (SNV).

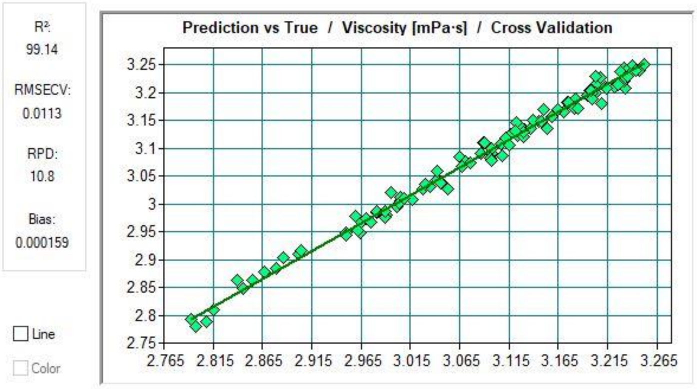
The predictive ability of this calibration method was tested as can be seen in Annex 1.

8.2.2. Viscosity calibration method

The calibration method performed to measure the viscosity prediction consisted of 119 samples but only 106 were used to build the calibration model. In the same way as in the acid value calibration model, the samples were introduced in duplicate so that a bigger library of spectra could be obtained.

The same procedure as the one performed for the acid value calibration model was followed. In this case, the spectral region used went from 9,403 to 4,397.4 cm^{-1} and all the optimization parameters were selected as can be seen in Figure 16.

In the same way as before, the number of samples was not big enough so a test set calibration could not be carried out, so a cross-validation was performed. The calibration line of this method can be observed in Figure 25.



R²: 99.14

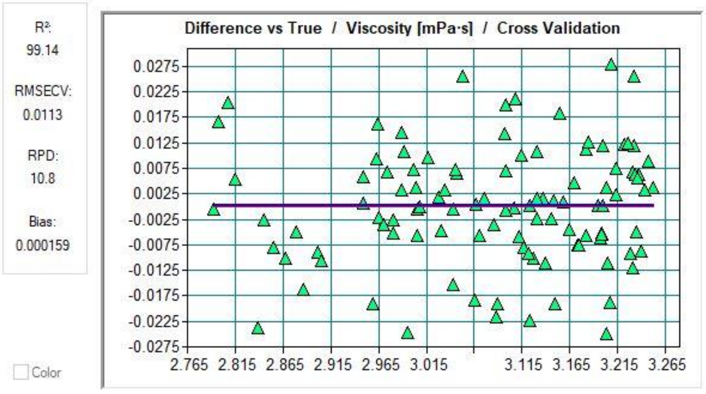
RMSECV: 0.0113

RPD: 10.8

Bias: 0.000159

Figure 25. Calibration line of the viscosity calibration method.

In the difference between the predicted value and the real value plot (Figure 26), in the same way as in the acid value method, the points are distributed in a homogeneous way above and below the line so it can be concluded that no tendency is found in the prediction of the model.


R²: 99.14

RMSECV: 0.0113

RPD: 10.8

Bias: 0.000159

☐ Color

Figure 26: Predicted value vs real value plot of the viscosity calibration model.

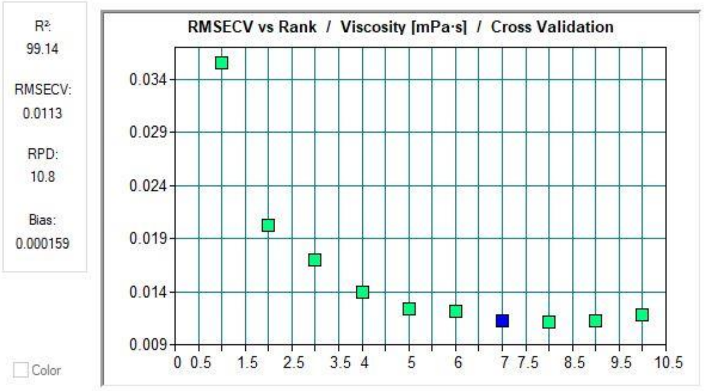
In Table 4, a summary of the calibration method can be observed.

Table 4: Results of the cross-validation in the acid value calibration model.

R^2	RMSECV	RPD	Bias	Rank	MD	Region	Pre-treatment	Calibration Samples	Total Samples
99,14	0,0113	10,8	0,000159	7	0,12	7501,8 - 6097,9 5025,7 - 4597,5	MSC	106	119

In the case of this calibration model, the Mahalanobis distance is higher than the acid value calibration model which means that this method is less restrictive than the other one. However, in both methods, the same number of outliers are eliminated by the software which means that those points are not representative of the methods.

In this method, the rank chosen is 7, which by looking at Figure 27, it can be concluded that is an appropriate number of factors as there is no underfitting nor overfitting.


R²: 99.14

RMSECV: 0.0113

RPD: 10.8

Bias: 0.000159

☐ Color

Figure 27: RMSECV vs Rank plot of the viscosity calibration method.

The pre-treatment the software has chosen when optimizing the method is Multiplicative Scattering Correction (MSC) and two spectral regions were chosen that corresponds to the 1st overtone of the oxygen-hydrogen bond and the combinations bands of the carbon-hydrogen bonds.

The predictive ability of this calibration method was tested as can be seen in Annex 2.

9. Conclusions

In this bachelor's thesis, it has been demonstrated that Near Infrared Spectroscopy is a suitable technique for the determination of chemical and physical properties in the esterification reaction of an acid and a glycol by means of Partial Least Squares model.

Near-Infrared Spectroscopy provides many advantages in the in-line monitoring of the manufacturing of polymeric resins. With this technique, information about the properties of the reaction is obtained in a fast way. In addition, the elimination of the need for sampling and the subsequent pre-treatment eases a lot the analysis. This implicates an increase in the plant personnel's safety as there is no need of extracting samples from the reactor and the reduction of waste.

It has been observed that factors such as the frequency region chosen, and the different spectral treatments affect the prediction results. As a consequence, the effect of these parameters on the RMSECV and the R^2 has been studied.

The influence of the number of factors used in the method has been studied too. From this, it can be concluded that the lower the number of factors, including all the spectral information of the property under study, the more accurate and precise the predicted value is, demonstrating the calibration model is able to filter all the undesired information.

All things considered, Near-Infrared Spectroscopy is a good technique to perform process monitoring in the chemical industry by means of the application of a proper Partial Least Squares multivariate calibration method.

10. Bibliography

(1) BASF https://www.basf.com/global/en/topics/diversity-inclusion.html (accessed Mar 2, 2021).

(2) Budde, F.; Krämer, K. Today's Chemical Industry: Which Way Is Up? *Value Creat.* **2003**, 1–9. https://doi.org/10.1002/3527600310.ch1.

(3) Castillo, M. Aplicación de La Espectroscopia NIR Al Control Analítico de Procesos de La Industria Química Tesis Doctoral, Universitat Autònoma de Barcelona, 2007.

(4) Thompson, J. M. Some Fundamentals of Infrared Spectroscopy. *Infrared Spectrosc.* **2018**, 1–33. https://doi.org/10.1201/9781351206037-1.

(5) Stretching-and-bending-vibrational-modes-The-and-signs-indicate-that-the-atoms_fig1_265226567 @ www.researchgate.net https://www.researchgate.net/figure/Stretching-and-bending-vibrational-modes-The-and-signs-indicate-that-the-atoms_fig1_265226567 (accessed Apr 7, 2021).

(6) Por, P.; Macias Macias Moisés Sabino Morales Espinosa Andrés Felipe León Rojas, K. Módelo Del Oscilador Armónico. Desarrollo Clásico, Mecanocuántico Y Aplicación En La Espectroscopia.

(7) Peguero, A. La Espectroscopia NIR En La Determinación de Propiedades Físicas y Composición Química de Intermedios de Producción y Productos Acabados., Universitat Autònoma de Barcelona, 2010.

(8) NIR Assignment Table. Bruker 2000, p 14286.

(9) Ringsted, T. Near Infrared Spectroscopy of Food Systems Using a Supercontinuum Laser, 2017.

(10) Blanco, M.; Villarroya, I. NIR Spectroscopy: A Rapid-Response Analytical Tool. *TrAC - Trends Anal. Chem.* **2002**, *21* (4), 240–250. https://doi.org/10.1016/S0165-9936(02)00404-1.

(11) Berard, C. W. Book Reviews: Book Reviews. In *Blood*; 1972; Vol. 40, p 612. https://doi.org/10.1016/S0006-4971(20)78019-5.

(12) Bruker Optik GmbH. *Multivariate Calibration.*; 2006.

(13) Blanco, M.; Cruz, J.; Armengol, M. Control Production of Polyester Resins by NIR Spectroscopy. *Microchem. J.* **2008**, *90* (2), 118–123. https://doi.org/10.1016/j.microc.2008.04.004.

(14) Azzouz, T.; Puigdoménech, A.; Aragay, M.; Tauler, R. Comparison between Different Data Pre-Treatment Methods in the Analysis of Forage Samples Using near-Infrared Diffuse Reflectance Spectroscopy and Partial Least-Squares Multivariate Calibration Method. *Anal. Chim. Acta* **2003**, *484* (1), 121–134. https://doi.org/10.1016/S0003-2670(03)00308-8.

(15) Kumar, S.; Andy, A. Fourier Transform-near Infrared Reflectance Spectroscopy Calibration Development for Screening of Oil Content of Intact Safflower Seeds. *Int. Food Res. J.* **2013**, *20* (2), 759–762.

(16) Bruker Optik GmbH. User Manual: Quant.; 2011

(17) Garcia, D. S. METODOLOGIA I VALIDACIÓ PER NIR, Universitat Rovira i Virgili, 2019.

(18) Montasell, T. Principios de OPUS / Quant Una Buena Calibración Comienza Antes de Utilizar El PLS.

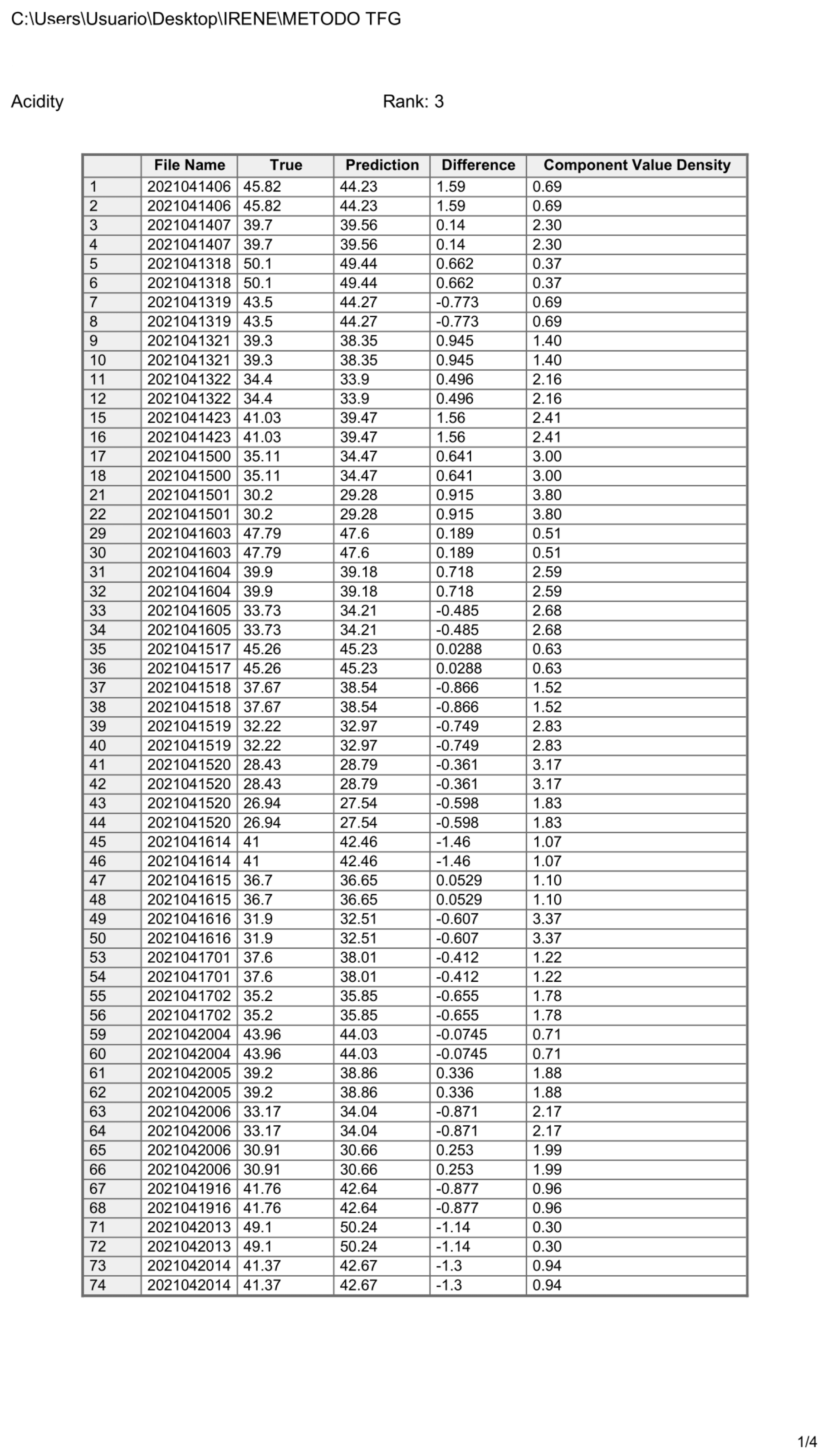
(19) matrix-f-ft-nir-spectrometer @ www.bruker.com https://www.bruker.com/en/products-and-solutions/infrared-and-raman/ft-nir-spectrometers/matrix-f-ft-nir-spectrometer.html (accessed Apr 21, 2021).

(20) process-reaction-monitoring @ www.bruker.com https://www.bruker.com/en/products-and-solutions/infrared-and-raman/opus-spectroscopy-software/process-reaction-monitoring.html (accessed Apr 21, 2021).

(21) base-package @ www.bruker.com https://www.bruker.com/en/products-and-solutions/infrared-and-raman/opus-spectroscopy-software/base-package.html (accessed Apr 21, 2021).

11. Annexes

Annexe 1: True vs Prediction - Acid value calibration model

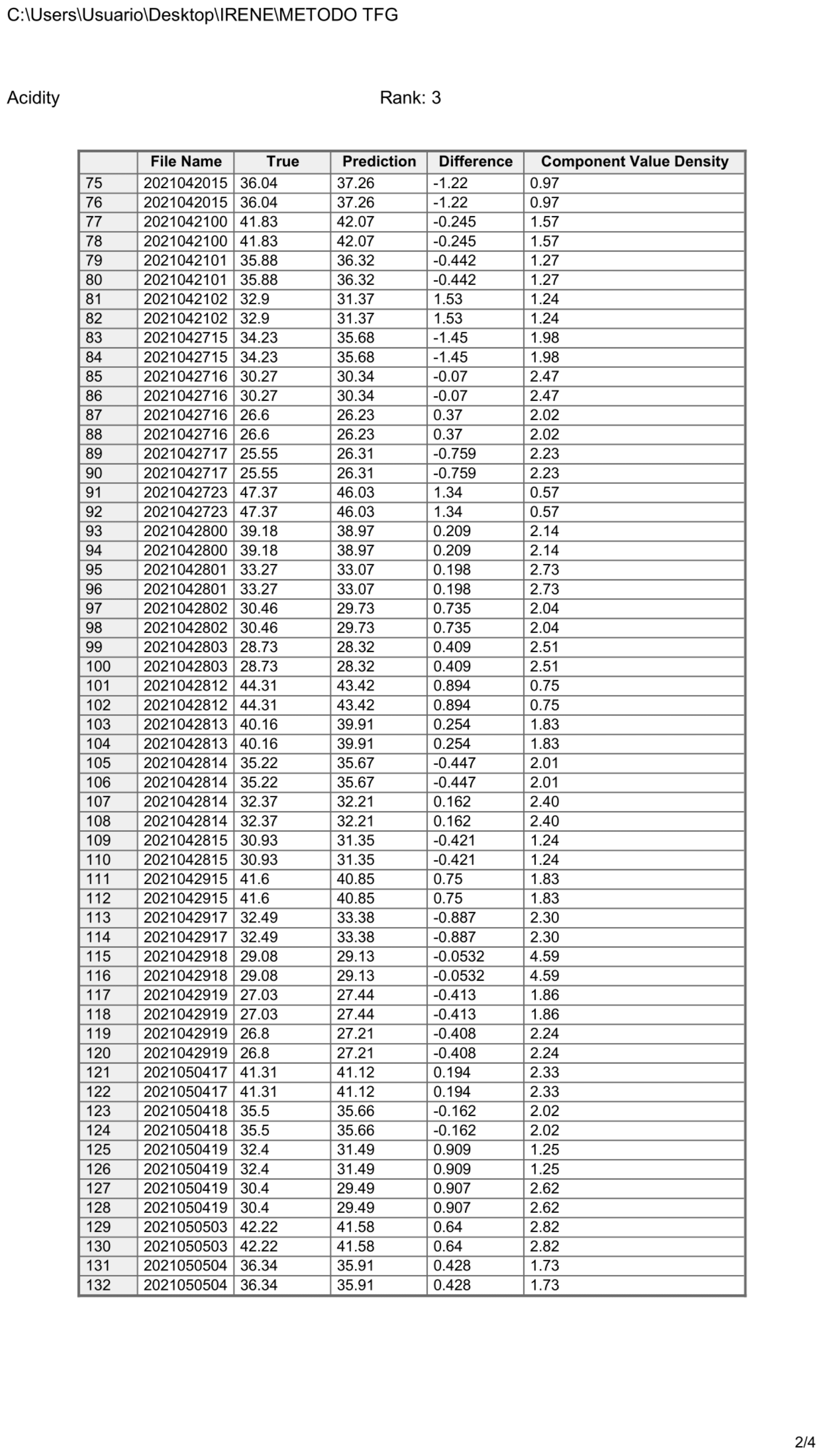


C:\Users\Usuario\Desktop\IRENE\METODO TFG

Acidity Rank: 3

	File Name	True	Prediction	Difference	Component Value Density
1	2021041406	45.82	44.23	1.59	0.69
2	2021041406	45.82	44.23	1.59	0.69
3	2021041407	39.7	39.56	0.14	2.30
4	2021041407	39.7	39.56	0.14	2.30
5	2021041318	50.1	49.44	0.662	0.37
6	2021041318	50.1	49.44	0.662	0.37
7	2021041319	43.5	44.27	-0.773	0.69
8	2021041319	43.5	44.27	-0.773	0.69
9	2021041321	39.3	38.35	0.945	1.40
10	2021041321	39.3	38.35	0.945	1.40
11	2021041322	34.4	33.9	0.496	2.16
12	2021041322	34.4	33.9	0.496	2.16
15	2021041423	41.03	39.47	1.56	2.41
16	2021041423	41.03	39.47	1.56	2.41
17	2021041500	35.11	34.47	0.641	3.00
18	2021041500	35.11	34.47	0.641	3.00
21	2021041501	30.2	29.28	0.915	3.80
22	2021041501	30.2	29.28	0.915	3.80
29	2021041603	47.79	47.6	0.189	0.51
30	2021041603	47.79	47.6	0.189	0.51
31	2021041604	39.9	39.18	0.718	2.59
32	2021041604	39.9	39.18	0.718	2.59
33	2021041605	33.73	34.21	-0.485	2.68
34	2021041605	33.73	34.21	-0.485	2.68
35	2021041517	45.26	45.23	0.0288	0.63
36	2021041517	45.26	45.23	0.0288	0.63
37	2021041518	37.67	38.54	-0.866	1.52
38	2021041518	37.67	38.54	-0.866	1.52
39	2021041519	32.22	32.97	-0.749	2.83
40	2021041519	32.22	32.97	-0.749	2.83
41	2021041520	28.43	28.79	-0.361	3.17
42	2021041520	28.43	28.79	-0.361	3.17
43	2021041520	26.94	27.54	-0.598	1.83
44	2021041520	26.94	27.54	-0.598	1.83
45	2021041614	41	42.46	-1.46	1.07
46	2021041614	41	42.46	-1.46	1.07
47	2021041615	36.7	36.65	0.0529	1.10
48	2021041615	36.7	36.65	0.0529	1.10
49	2021041616	31.9	32.51	-0.607	3.37
50	2021041616	31.9	32.51	-0.607	3.37
53	2021041701	37.6	38.01	-0.412	1.22
54	2021041701	37.6	38.01	-0.412	1.22
55	2021041702	35.2	35.85	-0.655	1.78
56	2021041702	35.2	35.85	-0.655	1.78
59	2021042004	43.96	44.03	-0.0745	0.71
60	2021042004	43.96	44.03	-0.0745	0.71
61	2021042005	39.2	38.86	0.336	1.88
62	2021042005	39.2	38.86	0.336	1.88
63	2021042006	33.17	34.04	-0.871	2.17
64	2021042006	33.17	34.04	-0.871	2.17
65	2021042006	30.91	30.66	0.253	1.99
66	2021042006	30.91	30.66	0.253	1.99
67	2021041916	41.76	42.64	-0.877	0.96
68	2021041916	41.76	42.64	-0.877	0.96
71	2021042013	49.1	50.24	-1.14	0.30
72	2021042013	49.1	50.24	-1.14	0.30
73	2021042014	41.37	42.67	-1.3	0.94
74	2021042014	41.37	42.67	-1.3	0.94

1/4

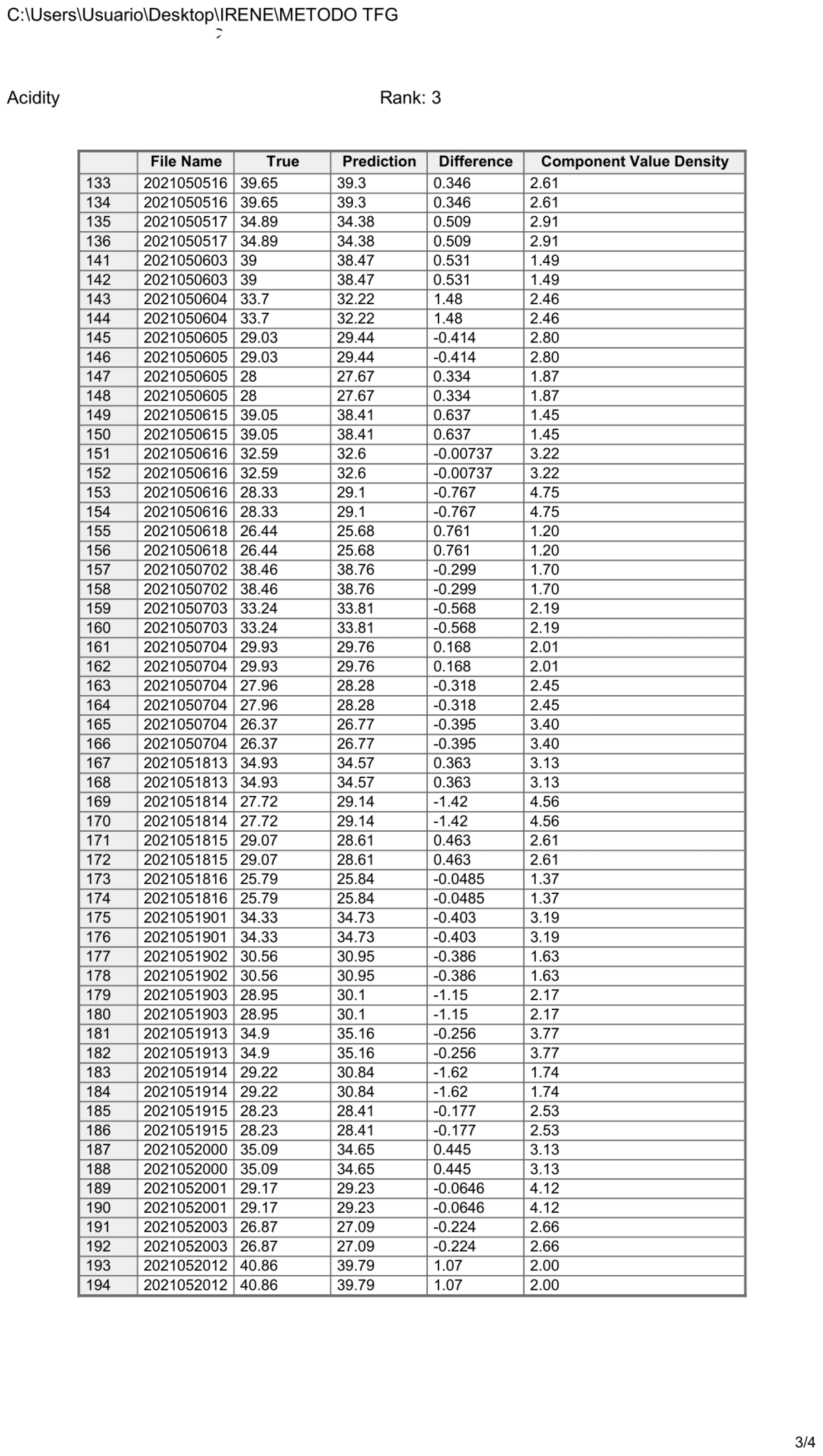


C:\Users\Usuario\Desktop\IRENE\METODO TFG

Acidity Rank: 3

	File Name	True	Prediction	Difference	Component Value Density
75	2021042015	36.04	37.26	-1.22	0.97
76	2021042015	36.04	37.26	-1.22	0.97
77	2021042100	41.83	42.07	-0.245	1.57
78	2021042100	41.83	42.07	-0.245	1.57
79	2021042101	35.88	36.32	-0.442	1.27
80	2021042101	35.88	36.32	-0.442	1.27
81	2021042102	32.9	31.37	1.53	1.24
82	2021042102	32.9	31.37	1.53	1.24
83	2021042715	34.23	35.68	-1.45	1.98
84	2021042715	34.23	35.68	-1.45	1.98
85	2021042716	30.27	30.34	-0.07	2.47
86	2021042716	30.27	30.34	-0.07	2.47
87	2021042716	26.6	26.23	0.37	2.02
88	2021042716	26.6	26.23	0.37	2.02
89	2021042717	25.55	26.31	-0.759	2.23
90	2021042717	25.55	26.31	-0.759	2.23
91	2021042723	47.37	46.03	1.34	0.57
92	2021042723	47.37	46.03	1.34	0.57
93	2021042800	39.18	38.97	0.209	2.14
94	2021042800	39.18	38.97	0.209	2.14
95	2021042801	33.27	33.07	0.198	2.73
96	2021042801	33.27	33.07	0.198	2.73
97	2021042802	30.46	29.73	0.735	2.04
98	2021042802	30.46	29.73	0.735	2.04
99	2021042803	28.73	28.32	0.409	2.51
100	2021042803	28.73	28.32	0.409	2.51
101	2021042812	44.31	43.42	0.894	0.75
102	2021042812	44.31	43.42	0.894	0.75
103	2021042813	40.16	39.91	0.254	1.83
104	2021042813	40.16	39.91	0.254	1.83
105	2021042814	35.22	35.67	-0.447	2.01
106	2021042814	35.22	35.67	-0.447	2.01
107	2021042814	32.37	32.21	0.162	2.40
108	2021042814	32.37	32.21	0.162	2.40
109	2021042815	30.93	31.35	-0.421	1.24
110	2021042815	30.93	31.35	-0.421	1.24
111	2021042915	41.6	40.85	0.75	1.83
112	2021042915	41.6	40.85	0.75	1.83
113	2021042917	32.49	33.38	-0.887	2.30
114	2021042917	32.49	33.38	-0.887	2.30
115	2021042918	29.08	29.13	-0.0532	4.59
116	2021042918	29.08	29.13	-0.0532	4.59
117	2021042919	27.03	27.44	-0.413	1.86
118	2021042919	27.03	27.44	-0.413	1.86
119	2021042919	26.8	27.21	-0.408	2.24
120	2021042919	26.8	27.21	-0.408	2.24
121	2021050417	41.31	41.12	0.194	2.33
122	2021050417	41.31	41.12	0.194	2.33
123	2021050418	35.5	35.66	-0.162	2.02
124	2021050418	35.5	35.66	-0.162	2.02
125	2021050419	32.4	31.49	0.909	1.25
126	2021050419	32.4	31.49	0.909	1.25
127	2021050419	30.4	29.49	0.907	2.62
128	2021050419	30.4	29.49	0.907	2.62
129	2021050503	42.22	41.58	0.64	2.82
130	2021050503	42.22	41.58	0.64	2.82
131	2021050504	36.34	35.91	0.428	1.73
132	2021050504	36.34	35.91	0.428	1.73

2/4

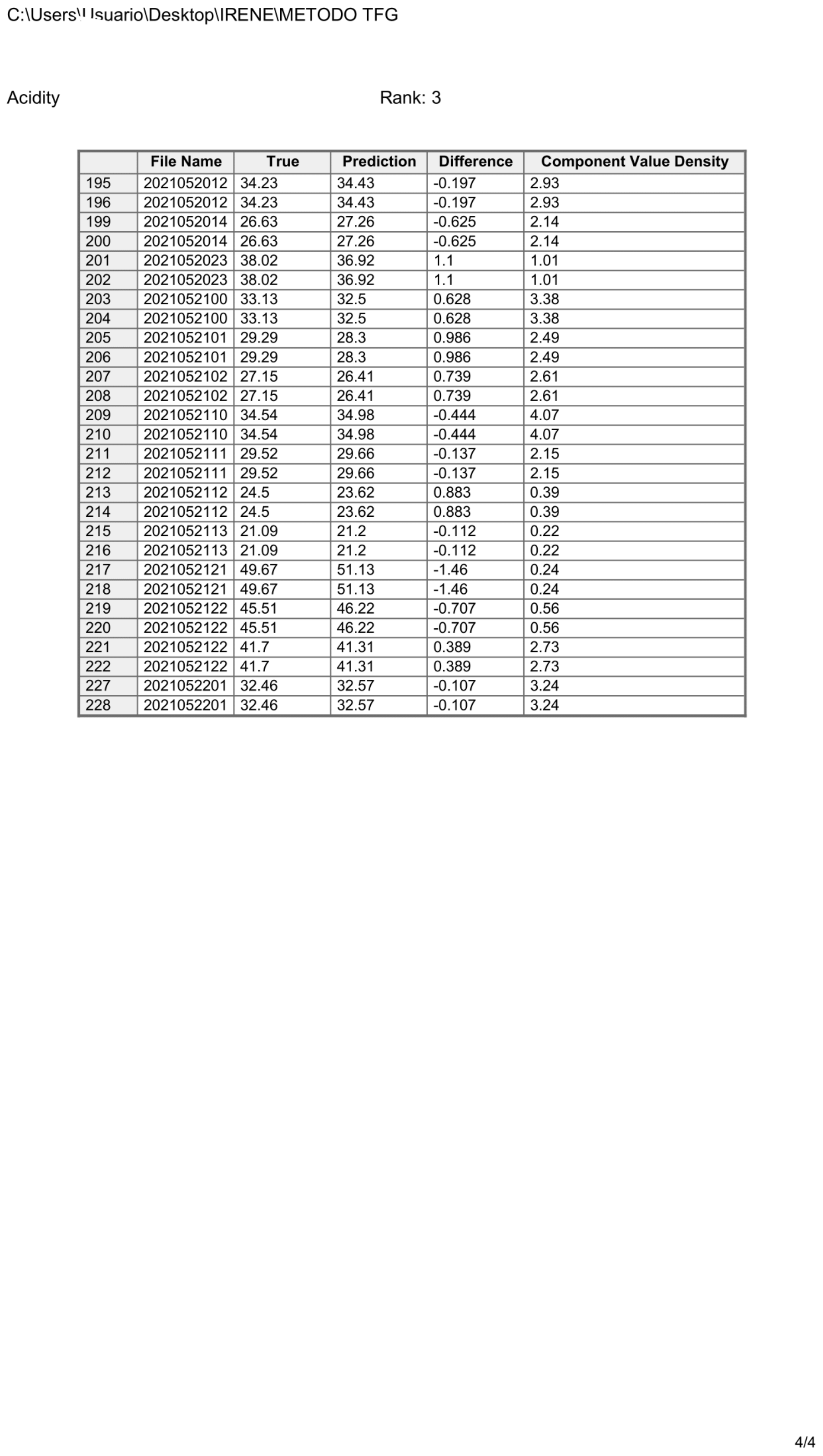


C:\Users\Usuario\Desktop\IRENE\METODO TFG

Acidity Rank: 3

	File Name	True	Prediction	Difference	Component Value Density
133	2021050516	39.65	39.3	0.346	2.61
134	2021050516	39.65	39.3	0.346	2.61
135	2021050517	34.89	34.38	0.509	2.91
136	2021050517	34.89	34.38	0.509	2.91
141	2021050603	39	38.47	0.531	1.49
142	2021050603	39	38.47	0.531	1.49
143	2021050604	33.7	32.22	1.48	2.46
144	2021050604	33.7	32.22	1.48	2.46
145	2021050605	29.03	29.44	-0.414	2.80
146	2021050605	29.03	29.44	-0.414	2.80
147	2021050605	28	27.67	0.334	1.87
148	2021050605	28	27.67	0.334	1.87
149	2021050615	39.05	38.41	0.637	1.45
150	2021050615	39.05	38.41	0.637	1.45
151	2021050616	32.59	32.6	-0.00737	3.22
152	2021050616	32.59	32.6	-0.00737	3.22
153	2021050616	28.33	29.1	-0.767	4.75
154	2021050616	28.33	29.1	-0.767	4.75
155	2021050618	26.44	25.68	0.761	1.20
156	2021050618	26.44	25.68	0.761	1.20
157	2021050702	38.46	38.76	-0.299	1.70
158	2021050702	38.46	38.76	-0.299	1.70
159	2021050703	33.24	33.81	-0.568	2.19
160	2021050703	33.24	33.81	-0.568	2.19
161	2021050704	29.93	29.76	0.168	2.01
162	2021050704	29.93	29.76	0.168	2.01
163	2021050704	27.96	28.28	-0.318	2.45
164	2021050704	27.96	28.28	-0.318	2.45
165	2021050704	26.37	26.77	-0.395	3.40
166	2021050704	26.37	26.77	-0.395	3.40
167	2021051813	34.93	34.57	0.363	3.13
168	2021051813	34.93	34.57	0.363	3.13
169	2021051814	27.72	29.14	-1.42	4.56
170	2021051814	27.72	29.14	-1.42	4.56
171	2021051815	29.07	28.61	0.463	2.61
172	2021051815	29.07	28.61	0.463	2.61
173	2021051816	25.79	25.84	-0.0485	1.37
174	2021051816	25.79	25.84	-0.0485	1.37
175	2021051901	34.33	34.73	-0.403	3.19
176	2021051901	34.33	34.73	-0.403	3.19
177	2021051902	30.56	30.95	-0.386	1.63
178	2021051902	30.56	30.95	-0.386	1.63
179	2021051903	28.95	30.1	-1.15	2.17
180	2021051903	28.95	30.1	-1.15	2.17
181	2021051913	34.9	35.16	-0.256	3.77
182	2021051913	34.9	35.16	-0.256	3.77
183	2021051914	29.22	30.84	-1.62	1.74
184	2021051914	29.22	30.84	-1.62	1.74
185	2021051915	28.23	28.41	-0.177	2.53
186	2021051915	28.23	28.41	-0.177	2.53
187	2021052000	35.09	34.65	0.445	3.13
188	2021052000	35.09	34.65	0.445	3.13
189	2021052001	29.17	29.23	-0.0646	4.12
190	2021052001	29.17	29.23	-0.0646	4.12
191	2021052003	26.87	27.09	-0.224	2.66
192	2021052003	26.87	27.09	-0.224	2.66
193	2021052012	40.86	39.79	1.07	2.00
194	2021052012	40.86	39.79	1.07	2.00

3/4



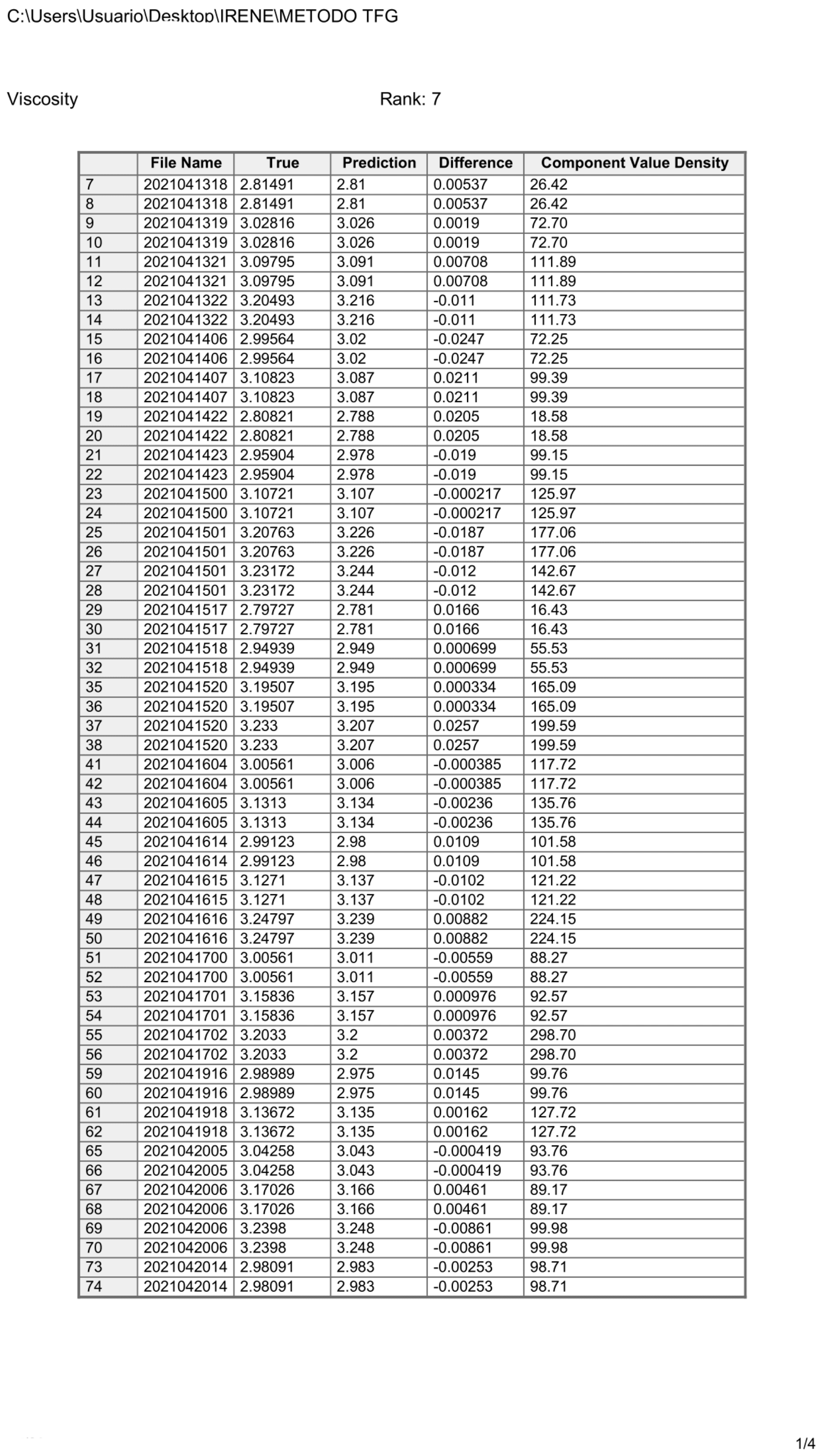
C:\Users\Usuario\Desktop\IRENE\METODO TFG

Acidity Rank: 3

	File Name	True	Prediction	Difference	Component Value Density
195	2021052012	34.23	34.43	-0.197	2.93
196	2021052012	34.23	34.43	-0.197	2.93
199	2021052014	26.63	27.26	-0.625	2.14
200	2021052014	26.63	27.26	-0.625	2.14
201	2021052023	38.02	36.92	1.1	1.01
202	2021052023	38.02	36.92	1.1	1.01
203	2021052100	33.13	32.5	0.628	3.38
204	2021052100	33.13	32.5	0.628	3.38
205	2021052101	29.29	28.3	0.986	2.49
206	2021052101	29.29	28.3	0.986	2.49
207	2021052102	27.15	26.41	0.739	2.61
208	2021052102	27.15	26.41	0.739	2.61
209	2021052110	34.54	34.98	-0.444	4.07
210	2021052110	34.54	34.98	-0.444	4.07
211	2021052111	29.52	29.66	-0.137	2.15
212	2021052111	29.52	29.66	-0.137	2.15
213	2021052112	24.5	23.62	0.883	0.39
214	2021052112	24.5	23.62	0.883	0.39
215	2021052113	21.09	21.2	-0.112	0.22
216	2021052113	21.09	21.2	-0.112	0.22
217	2021052121	49.67	51.13	-1.46	0.24
218	2021052121	49.67	51.13	-1.46	0.24
219	2021052122	45.51	46.22	-0.707	0.56
220	2021052122	45.51	46.22	-0.707	0.56
221	2021052122	41.7	41.31	0.389	2.73
222	2021052122	41.7	41.31	0.389	2.73
227	2021052201	32.46	32.57	-0.107	3.24
228	2021052201	32.46	32.57	-0.107	3.24

4/4

Annexe 2: True vs Prediction - Viscosity calibration model

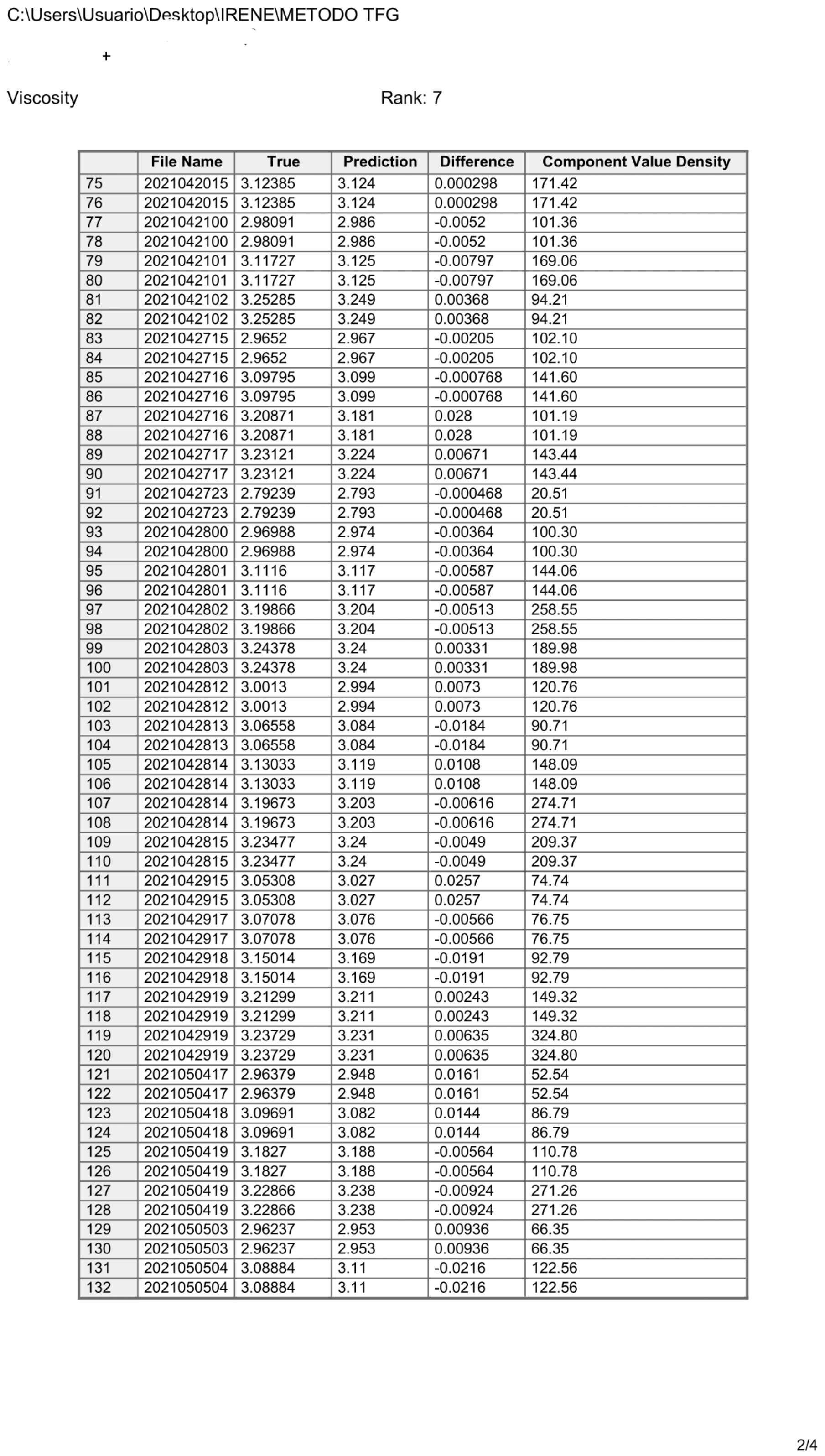


C:\Users\Usuario\Desktop\IRENE\METODO TFG

Viscosity Rank: 7

	File Name	True	Prediction	Difference	Component Value Density
7	2021041318	2.81491	2.81	0.00537	26.42
8	2021041318	2.81491	2.81	0.00537	26.42
9	2021041319	3.02816	3.026	0.0019	72.70
10	2021041319	3.02816	3.026	0.0019	72.70
11	2021041321	3.09795	3.091	0.00708	111.89
12	2021041321	3.09795	3.091	0.00708	111.89
13	2021041322	3.20493	3.216	-0.011	111.73
14	2021041322	3.20493	3.216	-0.011	111.73
15	2021041406	2.99564	3.02	-0.0247	72.25
16	2021041406	2.99564	3.02	-0.0247	72.25
17	2021041407	3.10823	3.087	0.0211	99.39
18	2021041407	3.10823	3.087	0.0211	99.39
19	2021041422	2.80821	2.788	0.0205	18.58
20	2021041422	2.80821	2.788	0.0205	18.58
21	2021041423	2.95904	2.978	-0.019	99.15
22	2021041423	2.95904	2.978	-0.019	99.15
23	2021041500	3.10721	3.107	-0.000217	125.97
24	2021041500	3.10721	3.107	-0.000217	125.97
25	2021041501	3.20763	3.226	-0.0187	177.06
26	2021041501	3.20763	3.226	-0.0187	177.06
27	2021041501	3.23172	3.244	-0.012	142.67
28	2021041501	3.23172	3.244	-0.012	142.67
29	2021041517	2.79727	2.781	0.0166	16.43
30	2021041517	2.79727	2.781	0.0166	16.43
31	2021041518	2.94939	2.949	0.000699	55.53
32	2021041518	2.94939	2.949	0.000699	55.53
35	2021041520	3.19507	3.195	0.000334	165.09
36	2021041520	3.19507	3.195	0.000334	165.09
37	2021041520	3.233	3.207	0.0257	199.59
38	2021041520	3.233	3.207	0.0257	199.59
41	2021041604	3.00561	3.006	-0.000385	117.72
42	2021041604	3.00561	3.006	-0.000385	117.72
43	2021041605	3.1313	3.134	-0.00236	135.76
44	2021041605	3.1313	3.134	-0.00236	135.76
45	2021041614	2.99123	2.98	0.0109	101.58
46	2021041614	2.99123	2.98	0.0109	101.58
47	2021041615	3.1271	3.137	-0.0102	121.22
48	2021041615	3.1271	3.137	-0.0102	121.22
49	2021041616	3.24797	3.239	0.00882	224.15
50	2021041616	3.24797	3.239	0.00882	224.15
51	2021041700	3.00561	3.011	-0.00559	88.27
52	2021041700	3.00561	3.011	-0.00559	88.27
53	2021041701	3.15836	3.157	0.000976	92.57
54	2021041701	3.15836	3.157	0.000976	92.57
55	2021041702	3.2033	3.2	0.00372	298.70
56	2021041702	3.2033	3.2	0.00372	298.70
59	2021041916	2.98989	2.975	0.0145	99.76
60	2021041916	2.98989	2.975	0.0145	99.76
61	2021041918	3.13672	3.135	0.00162	127.72
62	2021041918	3.13672	3.135	0.00162	127.72
65	2021042005	3.04258	3.043	-0.000419	93.76
66	2021042005	3.04258	3.043	-0.000419	93.76
67	2021042006	3.17026	3.166	0.00461	89.17
68	2021042006	3.17026	3.166	0.00461	89.17
69	2021042006	3.2398	3.248	-0.00861	99.98
70	2021042006	3.2398	3.248	-0.00861	99.98
73	2021042014	2.98091	2.983	-0.00253	98.71
74	2021042014	2.98091	2.983	-0.00253	98.71

1/4



C:\Users\Usuario\Desktop\IRENE\METODO TFG

+

Viscosity Rank: 7

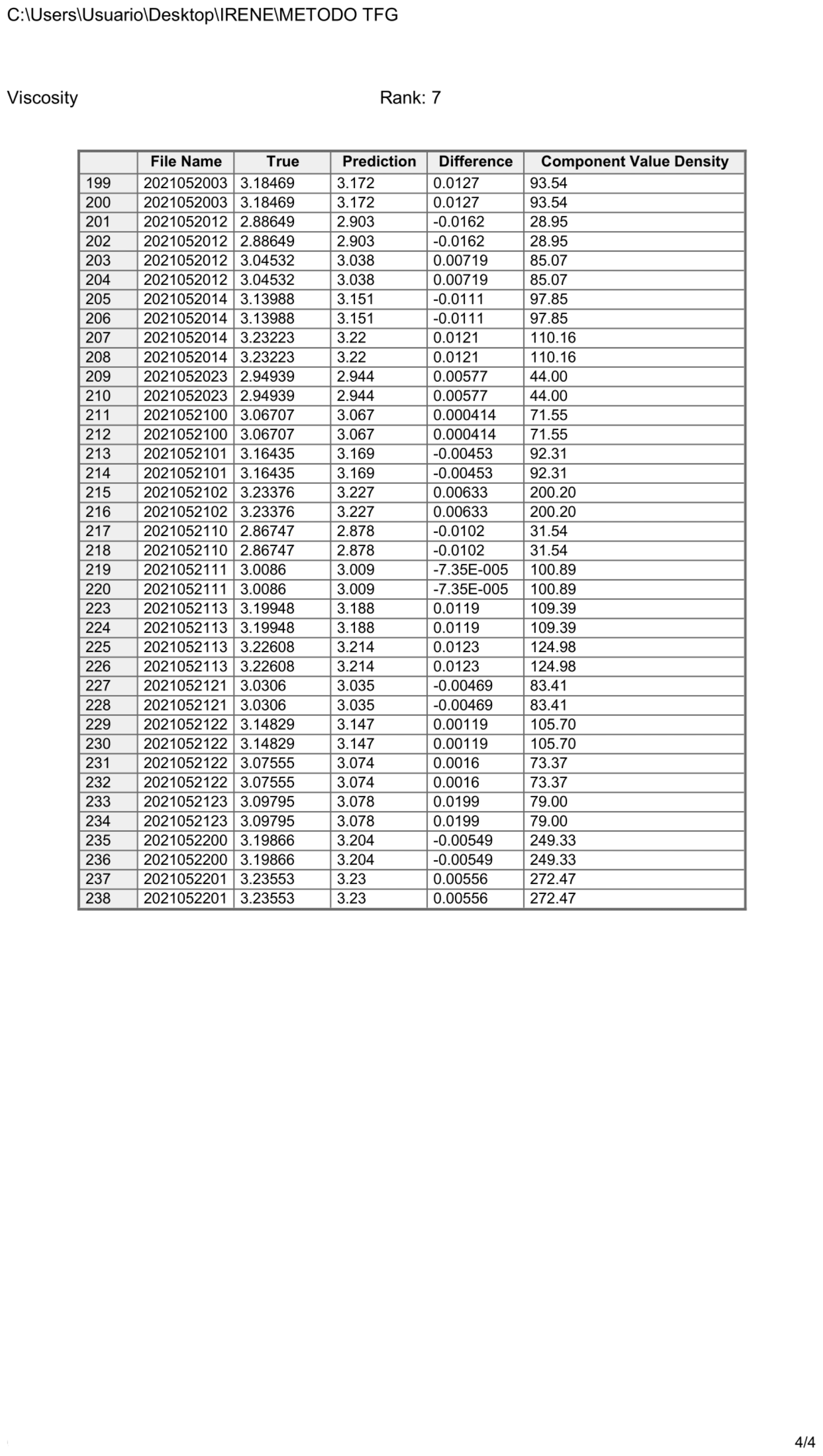
	File Name	True	Prediction	Difference	Component Value Density
75	2021042015	3.12385	3.124	0.000298	171.42
76	2021042015	3.12385	3.124	0.000298	171.42
77	2021042100	2.98091	2.986	-0.0052	101.36
78	2021042100	2.98091	2.986	-0.0052	101.36
79	2021042101	3.11727	3.125	-0.00797	169.06
80	2021042101	3.11727	3.125	-0.00797	169.06
81	2021042102	3.25285	3.249	0.00368	94.21
82	2021042102	3.25285	3.249	0.00368	94.21
83	2021042715	2.9652	2.967	-0.00205	102.10
84	2021042715	2.9652	2.967	-0.00205	102.10
85	2021042716	3.09795	3.099	-0.000768	141.60
86	2021042716	3.09795	3.099	-0.000768	141.60
87	2021042716	3.20871	3.181	0.028	101.19
88	2021042716	3.20871	3.181	0.028	101.19
89	2021042717	3.23121	3.224	0.00671	143.44
90	2021042717	3.23121	3.224	0.00671	143.44
91	2021042723	2.79239	2.793	-0.000468	20.51
92	2021042723	2.79239	2.793	-0.000468	20.51
93	2021042800	2.96988	2.974	-0.00364	100.30
94	2021042800	2.96988	2.974	-0.00364	100.30
95	2021042801	3.1116	3.117	-0.00587	144.06
96	2021042801	3.1116	3.117	-0.00587	144.06
97	2021042802	3.19866	3.204	-0.00513	258.55
98	2021042802	3.19866	3.204	-0.00513	258.55
99	2021042803	3.24378	3.24	0.00331	189.98
100	2021042803	3.24378	3.24	0.00331	189.98
101	2021042812	3.0013	2.994	0.0073	120.76
102	2021042812	3.0013	2.994	0.0073	120.76
103	2021042813	3.06558	3.084	-0.0184	90.71
104	2021042813	3.06558	3.084	-0.0184	90.71
105	2021042814	3.13033	3.119	0.0108	148.09
106	2021042814	3.13033	3.119	0.0108	148.09
107	2021042814	3.19673	3.203	-0.00616	274.71
108	2021042814	3.19673	3.203	-0.00616	274.71
109	2021042815	3.23477	3.24	-0.0049	209.37
110	2021042815	3.23477	3.24	-0.0049	209.37
111	2021042915	3.05308	3.027	0.0257	74.74
112	2021042915	3.05308	3.027	0.0257	74.74
113	2021042917	3.07078	3.076	-0.00566	76.75
114	2021042917	3.07078	3.076	-0.00566	76.75
115	2021042918	3.15014	3.169	-0.0191	92.79
116	2021042918	3.15014	3.169	-0.0191	92.79
117	2021042919	3.21299	3.211	0.00243	149.32
118	2021042919	3.21299	3.211	0.00243	149.32
119	2021042919	3.23729	3.231	0.00635	324.80
120	2021042919	3.23729	3.231	0.00635	324.80
121	2021050417	2.96379	2.948	0.0161	52.54
122	2021050417	2.96379	2.948	0.0161	52.54
123	2021050418	3.09691	3.082	0.0144	86.79
124	2021050418	3.09691	3.082	0.0144	86.79
125	2021050419	3.1827	3.188	-0.00564	110.78
126	2021050419	3.1827	3.188	-0.00564	110.78
127	2021050419	3.22866	3.238	-0.00924	271.26
128	2021050419	3.22866	3.238	-0.00924	271.26
129	2021050503	2.96237	2.953	0.00936	66.35
130	2021050503	2.96237	2.953	0.00936	66.35
131	2021050504	3.08884	3.11	-0.0216	122.56
132	2021050504	3.08884	3.11	-0.0216	122.56

2/4

C:\Users\Usuario\Desktop\IRENE\METODO TFG

Viscosity Rank: 7

	File Name	True	Prediction	Difference	Component Value Density
133	2021050516	3.00432	3.001	0.00376	124.23
134	2021050516	3.00432	3.001	0.00376	124.23
135	2021050517	3.13033	3.129	0.00157	167.14
136	2021050517	3.13033	3.129	0.00157	167.14
141	2021050603	2.90146	2.91	-0.00892	28.21
142	2021050603	2.90146	2.91	-0.00892	28.21
143	2021050604	3.04258	3.058	-0.0153	73.07
144	2021050604	3.04258	3.058	-0.0153	73.07
145	2021050605	3.14613	3.148	-0.0023	104.46
146	2021050605	3.14613	3.148	-0.0023	104.46
147	2021050605	3.17377	3.181	-0.00759	102.93
148	2021050605	3.17377	3.181	-0.00759	102.93
149	2021050615	2.90472	2.915	-0.0105	28.40
150	2021050615	2.90472	2.915	-0.0105	28.40
151	2021050616	3.03463	3.031	0.00329	80.47
152	2021050616	3.03463	3.031	0.00329	80.47
153	2021050616	3.12287	3.145	-0.0223	107.37
154	2021050616	3.12287	3.145	-0.0223	107.37
155	2021050618	3.23729	3.231	0.00636	324.10
156	2021050618	3.23729	3.231	0.00636	324.10
159	2021050703	3.0465	3.04	0.00646	88.50
160	2021050703	3.0465	3.04	0.00646	88.50
161	2021050704	3.15442	3.136	0.0184	124.55
162	2021050704	3.15442	3.136	0.0184	124.55
163	2021050704	3.18184	3.171	0.0112	93.86
164	2021050704	3.18184	3.171	0.0112	93.86
165	2021050704	3.22272	3.211	0.0121	149.11
166	2021050704	3.22272	3.211	0.0121	149.11
167	2021051813	2.8451	2.848	-0.00269	33.50
168	2021051813	2.8451	2.848	-0.00269	33.50
169	2021051814	2.97451	2.968	0.00688	101.53
170	2021051814	2.97451	2.968	0.00688	101.53
171	2021051815	3.08636	3.09	-0.00357	110.22
172	2021051815	3.08636	3.09	-0.00357	110.22
173	2021051816	3.17522	3.183	-0.00742	106.53
174	2021051816	3.17522	3.183	-0.00742	106.53
175	2021051817	3.21405	3.207	0.00742	206.17
176	2021051817	3.21405	3.207	0.00742	206.17
177	2021051901	2.83885	2.863	-0.0237	33.95
178	2021051901	2.83885	2.863	-0.0237	33.95
181	2021051903	3.08991	3.109	-0.0191	124.90
182	2021051903	3.08991	3.109	-0.0191	124.90
183	2021051904	3.2033	3.228	-0.0249	219.98
184	2021051904	3.2033	3.228	-0.0249	219.98
185	2021051913	2.8791	2.884	-0.00487	30.89
186	2021051913	2.8791	2.884	-0.00487	30.89
187	2021051914	3.01703	3.007	0.00969	108.72
188	2021051914	3.01703	3.007	0.00969	108.72
189	2021051915	3.12156	3.131	-0.00906	161.96
190	2021051915	3.12156	3.131	-0.00906	161.96
191	2021051916	3.20003	3.2	0.000168	299.88
192	2021051916	3.20003	3.2	0.000168	299.88
193	2021052000	2.85552	2.864	-0.00807	33.89
194	2021052000	2.85552	2.864	-0.00807	33.89
195	2021052001	2.98989	2.987	0.00322	102.75
196	2021052001	2.98989	2.987	0.00322	102.75
197	2021052003	3.11494	3.105	0.0102	121.33
198	2021052003	3.11494	3.105	0.0102	121.33



C:\Users\Usuario\Desktop\IRENE\METODO TFG

Viscosity Rank: 7

	File Name	True	Prediction	Difference	Component Value Density
199	2021052003	3.18469	3.172	0.0127	93.54
200	2021052003	3.18469	3.172	0.0127	93.54
201	2021052012	2.88649	2.903	-0.0162	28.95
202	2021052012	2.88649	2.903	-0.0162	28.95
203	2021052012	3.04532	3.038	0.00719	85.07
204	2021052012	3.04532	3.038	0.00719	85.07
205	2021052014	3.13988	3.151	-0.0111	97.85
206	2021052014	3.13988	3.151	-0.0111	97.85
207	2021052014	3.23223	3.22	0.0121	110.16
208	2021052014	3.23223	3.22	0.0121	110.16
209	2021052023	2.94939	2.944	0.00577	44.00
210	2021052023	2.94939	2.944	0.00577	44.00
211	2021052100	3.06707	3.067	0.000414	71.55
212	2021052100	3.06707	3.067	0.000414	71.55
213	2021052101	3.16435	3.169	-0.00453	92.31
214	2021052101	3.16435	3.169	-0.00453	92.31
215	2021052102	3.23376	3.227	0.00633	200.20
216	2021052102	3.23376	3.227	0.00633	200.20
217	2021052110	2.86747	2.878	-0.0102	31.54
218	2021052110	2.86747	2.878	-0.0102	31.54
219	2021052111	3.0086	3.009	-7.35E-005	100.89
220	2021052111	3.0086	3.009	-7.35E-005	100.89
223	2021052113	3.19948	3.188	0.0119	109.39
224	2021052113	3.19948	3.188	0.0119	109.39
225	2021052113	3.22608	3.214	0.0123	124.98
226	2021052113	3.22608	3.214	0.0123	124.98
227	2021052121	3.0306	3.035	-0.00469	83.41
228	2021052121	3.0306	3.035	-0.00469	83.41
229	2021052122	3.14829	3.147	0.00119	105.70
230	2021052122	3.14829	3.147	0.00119	105.70
231	2021052122	3.07555	3.074	0.0016	73.37
232	2021052122	3.07555	3.074	0.0016	73.37
233	2021052123	3.09795	3.078	0.0199	79.00
234	2021052123	3.09795	3.078	0.0199	79.00
235	2021052200	3.19866	3.204	-0.00549	249.33
236	2021052200	3.19866	3.204	-0.00549	249.33
237	2021052201	3.23553	3.23	0.00556	272.47
238	2021052201	3.23553	3.23	0.00556	272.47

4/4