THE APPLICATION OF EXPERIMENTAL DESIGN TO INVESTIGATE THE SOLVENT MATRIX EFFECTS OBSERVED DURING THE DETERMINATION OF RHODIUM (Rh) IN ORGANIC MEDIA BY GRAPHITE FURNACE ATOMIC ABSORPTION SPECTROMETRY (GFAAS)

by

ANTONIO BARATTA

submitted in accordance with the requirements for the degree of

MASTER OF SCIENCE

in the subject

CHEMISTRY

at the

UNIVERSITY OF SOUTH AFRICA

SUPERVISOR: PROF S O PAUL

CO-SUPERVISOR: DR J L FISCHER

NOVEMBER 2010

«The most exciting phrase to hear in science, the one that heralds new discoveries, is not “Eureka!” but “That's funny...” »

(Isaac Asimov)

Caro Niccolò, Scienza e Miracoli sono intorno a te...

Student number: **4549-141-0**

I declare that “The application of Experimental Design to investigate the solvent matrix effects observed during the determination of Rhodium (Rh) in organic media by Graphite Furnace Atomic Absorption Spectrometry (GFAAS)” is my own work and that all sources that I have used or quoted have been indicated and acknowledged by means of complete references.

_____ 15 December 2010

SIGNITURE DATE

(Mr A Baratta)

ACKNOWLEDGMENTS

- Johann Fischer for supporting this project.
- Prof Sylvia Paul for the availability and the advices.
- Harry Moloto for the SEM analyses.
- Sasol for this study opportunity.
- Lynette, grazie di tutto, soprattutto per Niccolò.
- Niccolò, grazie per questa nuova avventura.
- *Vi e' dell'oro, e delle perle assai; Ma le labbra di scienza sono un vaso prezioso. (Prov. 20.15).*

ABSTRACT

In an industrial application a GFAAS method for monitoring the Rh concentration in process streams is being used. Matrix effects are known to exist with the application of this technique; in fact, it was observed that different solvents lead to different results. Therefore, standard additions have to be employed for quantitative determinations, resulting in high costs and long analysis times. In an attempt to understand these interfering effects, fractional factorial designs were proposed to determine whether any GFAAS parameter was responsible for, or related to, the matrix effects. Seven GFAAS parameters were investigated: final temperature, ramp time and hold time of the transitions step (from the dry step); final temperature, ramp time and hold time of the ashing/pyrolysis step; ramp time of the atomisation step. The results showed that the matrix effects were not related to any specific parameter. A complete factorial design was implemented to demonstrate the fundamental role of the atomisation temperature. SEM analysis showed that the surface of the graphite tubes might be affected in different ways by different solvents. A Principal Component Analysis demonstrated that the matrix effects may be related to the viscosity and melting point of the solvents and may be independent of their molar mass. To identify the origins of these effects, an investigation on the link between the tube surface-sample matrix interactions and the physical properties of the matrices is recommended. Since GFAAS parameters cannot compensate for the matrix effects, standard additions remain the preferred mode of operation as it accounts for the effects *in-situ*.

CONTENTS

LIST OF ABBREVIATIONS

2-FD	2-Level Factorial Design
3-FD	3-Level Factorial Design
AAS	Atomic Absorption Spectrometry
Abs	Absorbance
AES	Atomic Emission Spectrometry
AFS	Atomic Fluorescence Spectrometry
ANOVA	Analysis Of Variance
BBD	Box-Bhenken Design
BP	Boiling Point
CA	Cluster Analysis
CCD	Central Composite Design
CFD	Complete Factorial Design
D	Density
DoE	Design of Experiment
FA	Factor Analysis
FAAS	Flame Atomic Absorption Spectrometry
FCC	Face Centred Cube
FEGSEM	Field Emission Gun Scanning Electron Microscope
FFD	Fractional Factorial Design
FIA	Flow Injection Analysis
GFAAS	Graphite Furnace Atomic Absorption Spectrometry
HCL	Hollow Cathode Lamp
Ht_{Ash}	Ashing Hold Time
Ht_{Atom}	Atomisation Hold Time
Ht_{Dry}	Drying Hold Time
Ht_{Tr}	Transition step Hold Time
ICP-MS	Inductively Coupled Plasma-Mass Spectrometry
ICP-OES	Inductively Coupled Plasma-Optical Emission Spectrometry
LDA	Linear Discriminant Analysis

MLR	Multiple Linear Regression
MM	Molar Mass
MP	Melting Point
N-Plot	Normal Probability Plot
OVAT	One Variable At a Time
PB	Plackett-Burman
PC	Principal Component
PCA	Principal Component Analysis
PCR	Principal Component Regression
PLS-R	Partial Least Squares Regression
ppb	Part per billion
ppm	Part per million
ppt	Part per trillion
ROPAC	Acetylacetonatocarbonyl(triphenylphosphine)rhodium(I)
RSD	Relative Standard Deviation
RSM	Response Surface Methodology
Rt_{Ash}	Ashing Ramp Time
Rt_{Atom}	Atomisation Ramp Time
Rt_{Dry}	Drying Ramp Time
Rt_{Tr}	Transition step Ramp Time
SEM	Scanning Electron Microscope
SIMCA	Soft Independent Modelling of Class Analogy
ST	Surface Tension
T_{Ash}	Ashing final Temperature
T_{Atom}	Atomisation final Temperature
T_{Dry}	Drying final Temperature
T_{Tr}	Transition step final Temperature
V	Viscosity
XRF	X-Ray Fluorescence

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

"There's two possible outcomes: if the result confirms the hypothesis, then you've made a discovery. If the result is contrary to the hypothesis, then you've made a discovery."[1]
(Enrico Fermi)

The hydroformylation reaction (or "oxo" synthesis) assisted by metal catalysis is a well known chemical and industrial process. In this context, Rh based catalysts are very popular as they provide milder reaction conditions and higher activity. One of the disadvantages of using Rh as a catalyst consists of its elevated price. Therefore, to minimise production costs, industrial units (plant sections) are normally supplied with Rh recovery procedures. Nevertheless, the separation of this precious metal from product streams is considerably challenging and under-recovery of Rh is a frequent and costly inconvenience.[2,3]

In order to minimise Rh losses, an analytical technique that can directly quantify this metal in industrial streams is required. The minimal loss of Rh can lead to a large financial impact, hence Rh should be detected at very low concentrations, such as *parts per billion* (ppb). Popular analytical techniques such as *Inductively Coupled Plasma-Optical Emission Spectrometry* (ICP-OES) and *X-Ray Fluorescence* (XRF) are not suitable to analyse elements at such low concentrations. *Graphite Furnace Atomic Absorption Spectrometry* (GFAAS), however, is a renowned technique capable of detecting and quantifying metals at concentrations as low as ppb and, in certain cases, ppt (*parts per trillion*).

1.1 Aim and scope of the project

A GFAAS method for quantifying low level Rh in aldehyde based matrices was developed in-house. During the method development studies, strong matrix effects were observed, where different solvents, containing the same amount of Rh, led to different results – i.e. different absorbance values (Figures 4.1 and 4.2). These matrix effects could not be compensated for by simply changing

the GFAAS operational parameters. Furthermore, due to the heterogeneous nature of the samples, matrices are never identical and the same external calibration (or even the same GFAAS method) cannot be used for all samples. Hence, the only alternative to yield reliable results was to adopt a standard addition approach, which involves higher costs and longer analysis times.

The basic premise of this project is to use Experimental Design to determine:

- whether different GFAAS operational parameters (factors/variables) have different effects when different solvents (matrices) are used; and
- whether by changing these parameters for different solvents (i.e. by having different furnace programmes), similar results can be obtained - like adjusting the car seat position for people of different heights.

A Fractional Factorial Design was employed to identify the important factors via the investigation of 5 different solvents: water, methanol, octanal, toluene and hexane. The results obtained from this approach dictated the subsequent steps of the project.

2 THEORETICAL BACKGROUND

"When you take stuff from one writer it's plagiarism; but when you take it from many writers, it's research."[4]
(Wilson Mizner)

Just 66 years have passed between the first flight of the Wright Brothers and the first step on the moon by Neil Armstrong. This illustrious example can be used to summarise the history of science. Since the scientific boom of the 20th century, one technological discovery followed another at a rapid and steady pace. For example, if one buys a computer today, by the next day that computer will already be outdated.

This rule is no exception for more fundamental sciences such as chemistry and physics. The main concern that scientists had in the beginning of the scientific revolution was *how to obtain data*. Parallel to the development of analytical instrumentations, the problem then evolved into *how to get more data* and *how to get better data*. Therefore, it should not be a surprise that having too many data is one of the main problems for scientists of today. To address this, the progress of technology has turned toward softer techniques that aim to support the existing tools. One of these new techniques is *chemometrics*, which was created by analytical scientists as an aid to analyse the large amount of data produced by the scientific community. Chemometrics is a collection of different tools. One of these tools is experimental design. This project is an attempt to conjugate experimental design with one of the most mature elemental analysis techniques, *viz*. GFAAS. Sections 2.1 to 2.3 provide a brief historical and theoretical background of these methodologies.

2.1 Spectrometry and Graphite Furnace Atomic Absorption

"Every great advance in science has issued from a new audacity of imagination."[5]
(John Dewey)

GFAAS is a spectroscopic technique, very useful for trace metal analysis as it is a quantification method based on mass. The principal difference between a GFAAS instrument and other atomic absorption spectrometers lies in the graphite tube/furnace, which can be considered the core of GFAAS (Figures 2.1 and 2.2).

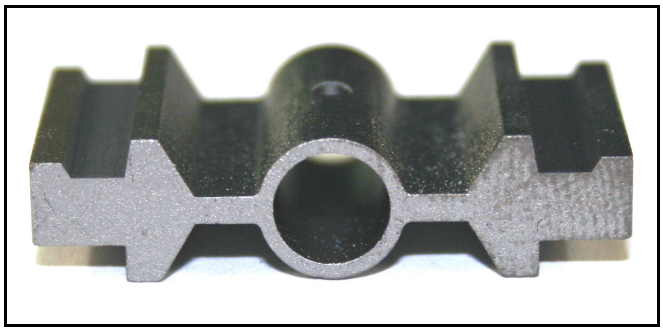

Figure 2.1: Front view of a graphite tube showing the atomisation cavity.

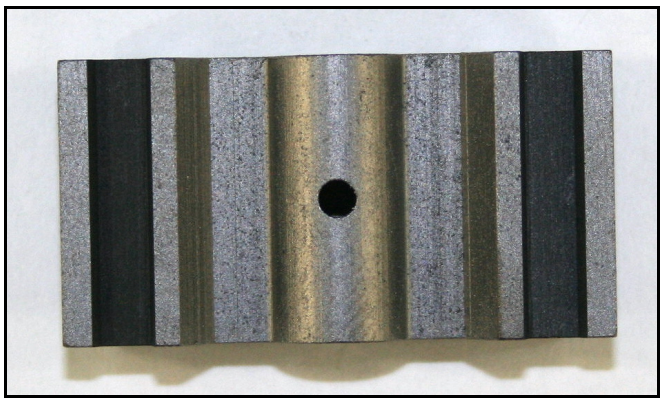

Figure 2.2: Top view of a graphite tube showing the injection orifice.

2.1.1 Spectroscopy and Spectrometry

The birth of spectroscopy can be traced to 1740 when Newton discovered that when white light passes through a prism its radiation splits into different colours. The first atomic spectroscopic device was created in 1856 when Bunsen

introduced the Bunsen burner. This can be considered as the first Atomic Emission Spectrometry (AES) instrument. This instrument demonstrated that the lines of a spectrum originate from elements and not from compounds (molecules). With the development of more sophisticated flame emission devices a few years later, Rb, Cs, Tl and In were discovered.[6,7]

The term *spectrometry* refers to the techniques/instruments that use spectroscopic principles for qualitative and quantitative analysis. Hence, the introduction of the Bunsen burner makes 1856 the birth year of atomic spectrometry.

There are three groups of atomic spectrometric techniques:[8]

- *optical spectrometry*: based on the intensity of absorbed or emitted light/energy;
- *mass spectrometry*: based on the nature and amount of ions generated from a sample; and
- *electron spectrometry*: based on the kinetic energy of the electrons ejected from a sample (commonly referred to as *electron spectroscopy*).

Optical atomic spectrometry is based on three mechanisms (these descriptions deal with inorganic spectroscopy):[7,8]

- *absorption* of light/energy by free atoms in the gas phase: Atomic Absorption Spectrometry (AAS);
- *emission* of photons/energy by free atoms/ions when a sample is excited with thermal or electric energy: Optical Emission Spectrometry (OES); and
- *fluorescence,* which is a stimulated emission of energy when a sample is excited with radiation energy (i.e. absorption followed by emission): Atomic Fluorescence Spectrometry (AFS).

Absorption occurs when an electron undergoes an electronic transition from a ground energy state to an excited energy state. Emission occurs when an electron returns from an excited state to a ground state. Fluorescence is a combination of these two processes: an electron initially absorbs energy and

moves to an excited energy state; then it emits energy while it returns to the ground state. Figure 2.3 is a simple schematic representation of these mechanisms. Light is absorbed or emitted at a precise wavelength, which is specific for each element. Knowing this wavelength allows the identification of unknown elements – qualitative analysis. The quantification is related to the intensity of the absorbed/emitted energy.[6,8]

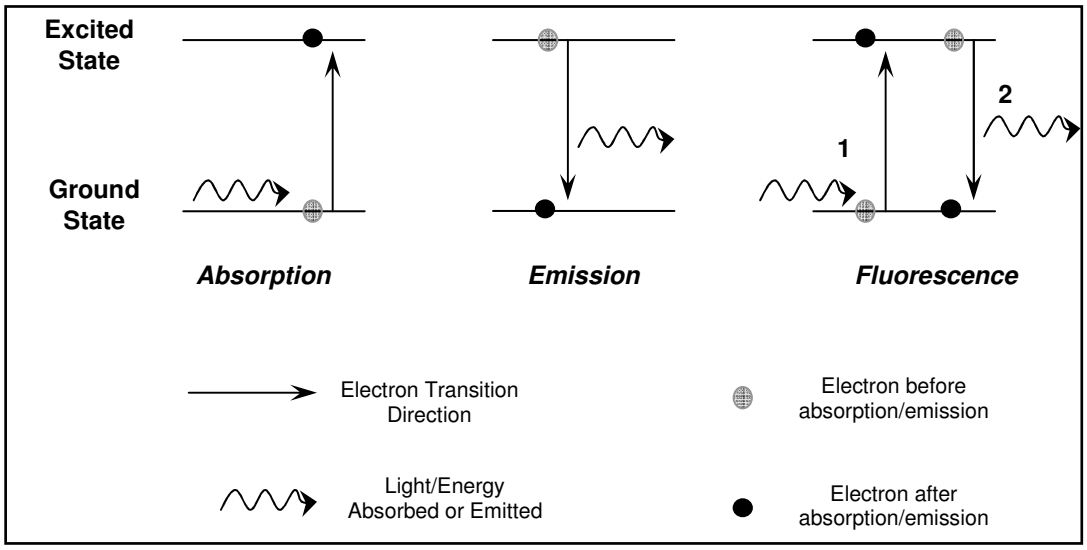

Figure 2.3: Schematic representation of the different optical spectroscopy mechanisms.

GFAAS, as the name implies, is an atomic absorption technique. When a beam of light passes through a sample, the atomised atoms will absorb a fraction of this light. The fraction of light that it is not absorbed, and passes through the atoms, is associated with a quantity called *transmittance* (T):

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

where: T is the transmittance; I_0 is the intensity of the incident light (from the source); I is the intensity of the light fraction that was not absorbed (by the sample).

AAS is mainly used for quantitative analysis. The quantification (at a given wavelength) is based on the Beer-Lambert Law:

$$A = -\log T = \log \frac{I}{I_0} = abc$$

where: A is the *absorbance*; a is the absorption coefficient; b is the sample path length; and c is the analyte concentration. The absorbance (A) is measured by the instrument; a and b are constants. Consequently, via this equation, one can obtain the concentration (c) of unknown samples.[6-8]

2.1.2 Graphite Furnace Atomic Absorption Spectrometry

In 1959 Boris L'vov introduced the idea of using a graphite furnace tube as a reactor for the generation of free atoms, i.e. as an atomiser. Since then, GFAAS went through many modifications and developments with the contribution of L'vov himself.[7,9]

The first commercial GFAAS was introduced in 1970 by Perkin-Elmer. The introduction of GFAAS into the market was rather successful. This success was mainly due to the total atomisation achievable in a tube atomiser.[10] In the subsequent years several studies were conducted to understand and optimise the atomisation process and improve the graphite tube performance.[9,11-15]

A typical set-up of an AAS instrument is shown in the block diagram in Figure 2.4.[6]

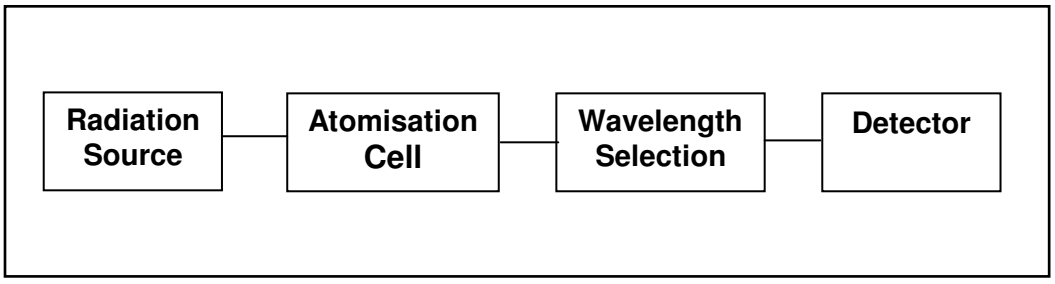

Figure 2.4: Block Diagram of an AAS instrument.[6]

In the case of GFAAS, the radiation source is a Hollow-Cathode Lamp (HCL), the atomisation cell is the graphite tube and the wavelength is selected via a monochromator (Figure 2.5).

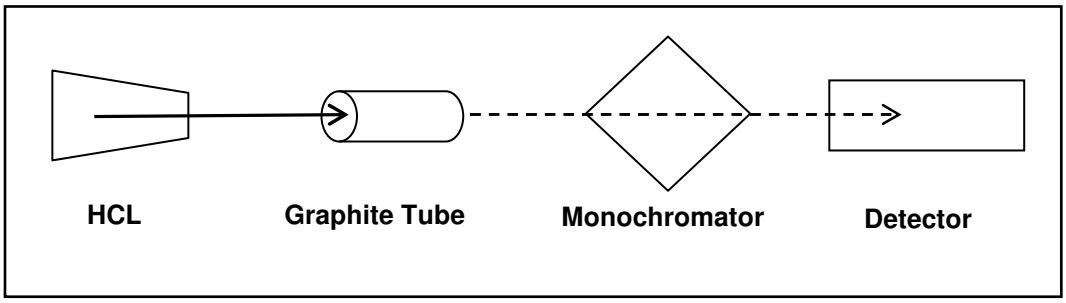

Figure 2.5: Representation of a GFAAS set up. The solid arrow represents the light before passing through the sample (I_0). The dotted arrow represents the non-absorbed fraction of light (I).

A GFAAS analysis is mainly characterised by its furnace programme which dictates what occurs inside the tube. This programme is composed of four key steps (Figure 2.6):

a. *Injection*
b. *Drying (to dispose of the solvent)*
c. *Ashing/Pyrolysis (to dispose of the matrix)*
d. *Atomisation*

The injection of the sample is done by a probe that takes the sample from a vial to the tube (Figure 2.6a). This can be done at room temperature or after having pre-heated the tube to 40-50 ºC.

During the drying step the temperature is increased slowly so that the evaporation of the solvent occurs under controlled and mild conditions (Figure 2.6b). Normally, the final temperature is slightly below or above the boiling point of the solvent. It is very important that this step is not too rapid. If the solvent splashes (known as *sputtering*) the analyte will not be distributed evenly on the tube surface and some atoms might be carried out of the tube by the gas flow. This affects the repeatability and the accuracy of the analysis.[7] Once the drying step is completed, an invisible layer of organic matrix remains on the tube surface (Figure 2.6b).

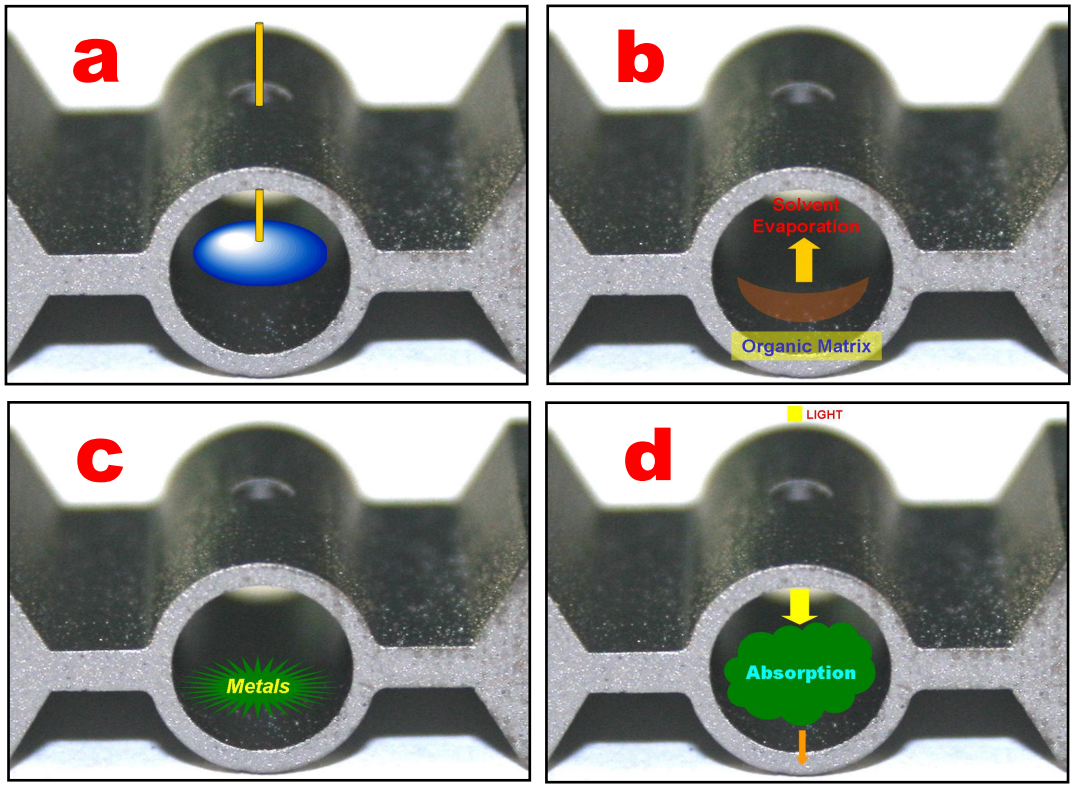

Figure 2.6: Furnace program steps: a) Injection; b) Drying step (the orange semicircle represents an invisible organic layer); c) end of Ashing Step (metal oxides and carbides remain on the tube surface); d) Atomisation and Absorption Step (the orange arrow represents the non-absorbed fraction of light).

The purpose of the ashing step is to destroy the organic matrix by elevating the temperature. This temperature must be lower than the atomisation temperature of the analyte to avoid sample loss – typically, it is in the range 400 to 1000 ºC. Once the organic matter has been pyrolysed and removed by the gas flow, the analyte, in the form of metal oxides or carbides, should be the only matter remaining in the tube (Figure 2.6c). The ashing step is critical when the analyte is dispersed in a mixture of different organic components. The organic matrix must be totally destroyed to avoid interferences during atomisation. However, an analyst must be careful to avoid analyte losses while removing the organic matter.[7]

The elements that remain on the surface of the tube must be atomised to absorb light. During the atomisation step the temperature is rapidly (1-2 seconds) raised

to the atomisation temperature of the element (each element is characterised by a specific atomisation temperature).[6] During the atomisation, an atomic cloud of free atoms is formed and a beam of light from the HCL is sent through the tube (Figure 2.6d). The light is absorbed by the free atoms and an absorption peak/signal is generated.

An example of a furnace program is shown in Table 2.1. Every element requires a different furnace program, which also depends on the matrix. If one considers every element/matrix combination as a different meal, the furnace program will correspond to the relevant recipe.

Table 2.1: Example of a furnace programme.

Step	***Final Temp (ºC)***	***Ramp Time (s)***	***Hold Time (s)***	***Inert gas***
Pre-heating	40	1.0	2.0	On
Injection				On
Drying	80	20	20	On
Transition	150	10	10	On
Ashing	800	20	10	On
Stabilisation	800	0.1	1.0	Off
Atomisation (signal reading)	2400	1.0	1.0	Off
Cleaning	2600	1.0	1.0	On

GFAAS is a technique with high sensitivity and low detection limits (10-100 ppb). One advantage of this technique when compared to other spectrometric techniques consist of its low capital and maintenance costs; another advantage can be identified in the negligible sample amount consumed during the analysis.[6,7]

GFAAS users have to face two major problems: slowness of analysis (only one element can be analysed at a time) and matrix interferences (especially with organic based samples).[6,7,16,17] Nevertheless, GFAAS has often been preferred to ICP-OES and Flame AAS (FAAS) because of its lower detection limits and higher sensitivity.

Since the commercialisation of ICP-Mass Spectrometry (ICP-MS), GFAAS is facing strong competition because ICP-MS combines all the advantages of the other spectrometers (fast; very low detection limits; robust; multi element analysis; simple spectra; wide working linear range). On the other hand, the capital and maintenance costs of ICP-MS instruments are very high and, with a limited budget, GFAAS is often a better alternative.[6,16]

ICP-MS is the elemental analysis technique of today and it has also been used together with chemometric techniques.[18-20] Nonetheless, research on and via GFAAS is still active today,[21,22] long after the first articles on graphite furnace utilisation were published (well before its commercialisation).[9,23,24] The principal application of GFAAS is the detection and quantification of trace metals. This is needed in different fields, such as petrochemical, medical and environmental applications.[21,25-30]

2.2 A few words about Chemometrics

"Data is not information, information is not knowledge, knowledge is not understanding, understanding is not wisdom."[31]
(Clifford Stoll)

"Just the mere mention of the word chemometrics invokes images of incomprehensible statistics and mathematics that are more black magic than a useful tool."[32] This was written by James Duckworth in 1998 in a paragraph entitled "What is 'Chemometrics'?". From this statement, one might think that it should be easier to explain what chemometrics is not, rather than what it is.

Think of the chemometric tool box as a Russian *Matryoshka* doll (Figure 2.7). Inside this chemometric tool box there are other toolboxes. The main difference is that chemometrics is not *univariate* (only one doll inside another) but *multivariate*; where inside a bigger doll there might be more than one smaller doll, which in turn might hold other smaller dolls.

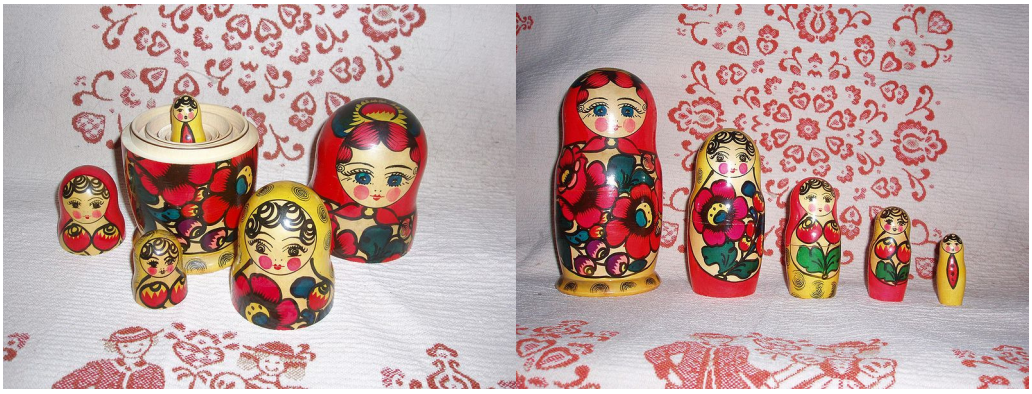

***Figure 2.7: Russian Matryoshka doll.*[33]**

One of the main difficulties with chemometrics, is to understand which tool box is the most suitable for a particular problem[32] – *when to use it* can be more important than *how to use it*.

The need for chemometrics arose in analytical chemistry when the translation from analytical data (numbers) to useful information was becoming uncertain because of large amount of data. The solution to this was the chemometric *multivariate analysis*, sparked by the development of computers.[34-37]

It is commonly accepted that the birth of chemometrics occurred in 1971, when Prof. Svante Wold coined the Swedish term “kemometri”, which became *chemometrics* in English (chemometrics is the *-metrics* of chemistry).[38-40] However, Svante Wold stated that chemometrics was actually born around 1905, when Gosset, a chemist, invented the *t-test* to include statistics in chemistry.[34,35] According to *The International Chemometrics Society*, chemometrics is "the chemical discipline that uses mathematical and statistical methods to, (a) design or select optimal measurement procedures and experiments; and (b) provide maximum chemical information by analysing chemical data”.[39,41]

Svante Wolde and Bruce R. Kowalski founded *The International Chemometrics Society* on June 10, 1974. Chemometrics officially started its journey on this date and Wolde and Kowalski are considered the fathers of chemometrics.[34,40,42-44] Kowalski published the first article and the first book with the term *chemometrics*

in the title[35,42] and issued the first Chemometrics Newsletter.[35,38] Kowalski also simplified the definition of chemometrics by saying that every method that is used to extract useful *chemical* information from raw data can be considered a chemometric tool.[42] But the simplest definition of chemometrics that can be found in literature is probably the following: "chemometrics is what chemometricians do".[45]

Two very important events in the history of chemometrics were the founding of "Chemometrics and Intelligent laboratory Systems" (1986) and the "Journal of Chemometrics" (1987).[35,38] Nevertheless, for the popularity of this discipline, the teaching of chemometric techniques as part of the analytical chemistry curriculum was essential.[46] Before chemometrics, analytical chemistry was considered a science without a theory.[47] The World-Wide Web also contributed to expand the chemometrics community, especially when some websites begun to offer free, simple chemometrics software.[48,49] On the other hand, since the free distribution of chemometric software, chemometrics has been used, and misused, also by users without any training.[39]

Chemometrics can be applied to every scientific field;[40,50,51] so it is almost impossible to give an example for every chemometric application. The petrochemical industry is an example of a field that is being increasingly assisted by chemometric techniques.[52,53]

GFAAS has often been coupled with chemometrics for different goals, such as agricultural, environmental and petrochemical applications.[25,27,54,55] On the other side of this medal, chemometric tools have been implemented to solve GFAAS problems such as interferences and matrix effects.[12,56-62]

As mentioned above, chemometrics encloses a series of tools that can be used to solve different problems. Among these tools, the most widespread ones are *Principal Component Analysis* (PCA), *Multiple Linear Regression* (MLR), *Principal Component Regression* (PCR), *Partial Least Squares Regression* (PLS-R) and *Experimental Design* (or *Design of Experiment*, DoE).[63] Sections 2.2.1 to

2.2.4 give a brief and simple illustration of these techniques. Experimental Design is dealt with in more detail in Section 2.3.

2.2.1 Principal Component Analysis

PCA is probably the most common exploratory (and pattern recognition) method in chemometrics, well represented in literature.[37,51-53,55,64] The main outcome of a PCA is the reduction of the *dimensionality* of a data set, to facilitate the interpretation of the data without losing any significant information.[8,52,63,65] The dimensionality is proportional to the amount of variables, while the human mind can only visualise and interpret data in a 2- or 3-dimensional setting. The dimensionality reduction is achieved by replacing the original variables with a reduced number of uncorrelated variables called the *principal components* (PCs).[53,63]

When the PCs are generated, most of the information hidden in the data set is allocated to the first PC (PC1). The second PC (PC2) represents the maximum variance of the knowledge not explained in PC1. PC2 is uncorrelated and *orthogonal* to PC1. PC3 is orthogonal to and explains the variance not included in PC1 and PC2. The same reasoning holds for PC4 and so forth. The higher the number of PCs, the more complicated the interpretation of the results. Ideally, each PC should represent a particular hidden phenomenon and/or structure. The aim is to have the smallest number possible of PCs (optimum PC number, *A*) without losing any significant information. Generally, the higher PCs do not contain valuable information, but only the background noise of the analysis.[53,63]

The PCs are used to construct a new system of coordinates where the PCs are the axes. Then, the data matrix (composed of number of objects/samples and variables) is decomposed into *score* and *loading* vectors to form new matrices. The *score* vectors represent the samples while the *loading* vectors represent the variables. *Scores* and *loadings* are the projections (and the coordinates) of each object and variable onto the PCs. As a consequence, *score* and *loading* plots are generated. The *score* plots describe the relationships between each sample;

the *loading* plots are an indication of the importance of (i.e. how large is the influence) and the relations between the different variables.[52,63,64]

These plots are used to inspect the data set. Using two or three PCs at a time, the data set can be arranged in 2- or 3-dimensions and an analyst can study how the information is distributed among the different PCs. PC1 explains most of the knowledge and it is typically set as the *x*-axis of the *score* and *loading* plots. They are very useful for the identification of clusters or classes and of the most influential variables.[63]

Figure 2.8 is a representation of a of *score* plot (adapted from Brereton, 2007).[64] In this example, various elements are plotted in relation to their physical constants (see Figure 2.9). *Score* plots help to discern which elements have the same characteristics according to the measured variables. In Figure 2.8 one can distinguish three distinct groups: halogens, inert gases and metals. In general, related objects have similar scores and can be grouped together in a specific section of the plot (i.e. they are close to each other). Objects that are close to the origin of the axes are the most common ones and do not present any "abnormality".[53,63,64]

Figure 2.9 shows the corresponding *loading* plot for the *score* plot in Figure 2.8 (adapted from Brereton, 2007).[64] Melting point and boiling point are very close to each other (they show a strong positive correlation) which means that they measure and indicate something similar. Also density and oxidation number indicate something which is related to melting and boiling point (density, melting point and boiling point depend on temperature). On the other hand, electronegativity is almost at 90º from the other variables, which indicates that there is no correlation between them. In general, important variables are far from the axis origin. Variables close to each other and far from the origin have strong positive covariance (they produce strong, similar effects). Variables far from the origin from opposite sides (roughly along a straight line) have strong negative covariance, producing strong but opposite effects. Variables that are at 90º are independent from each other and variables that are close to the origin have a small or null importance.[63,64]

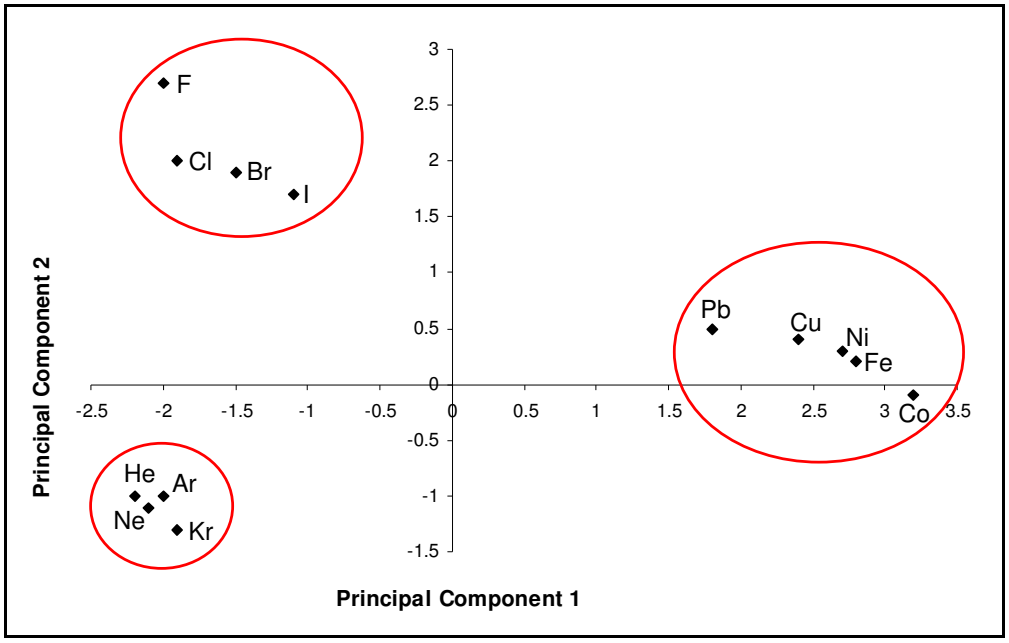

Figure 2.8: Example of a score plot, with PC1 on the x-axis and PC2 on the y-axis. Three classes of elements can be distinguished: halogens, inert gases and metals (adapted from Brereton, 2007).[64]

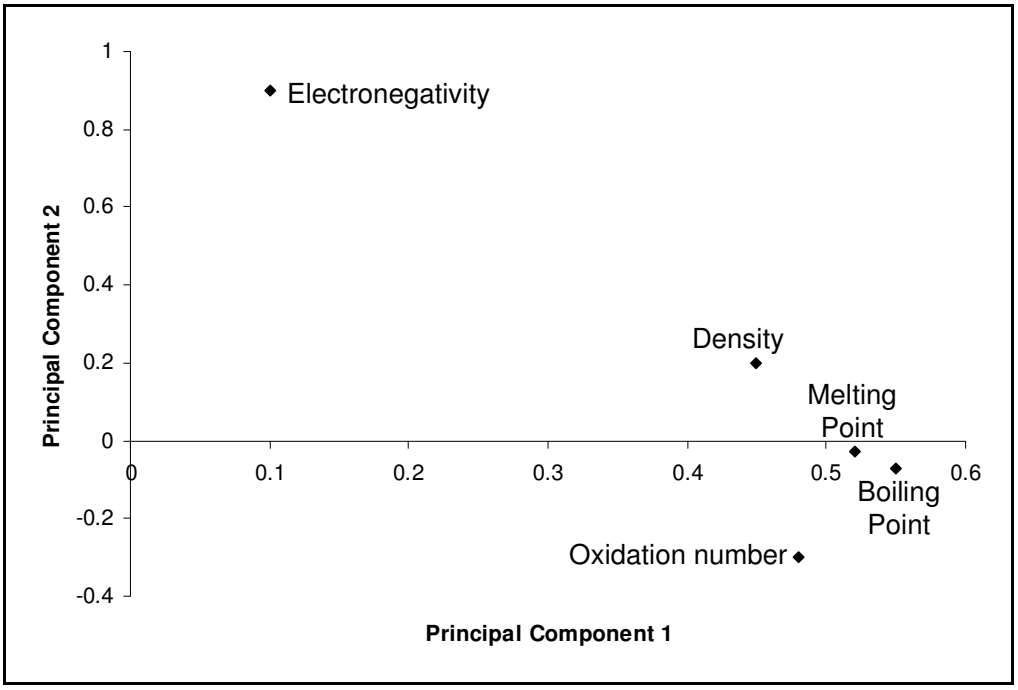

Figure 2.9: Example of a loading plot, with PC1 on the x-axis and PC2 on the y-axis. Melting point and boiling point show a strong positive co-variance; electronegativity is independent of the other variables (adapted from Brereton, 2007).[64]

An important step in PCA is the direct comparison between *score* and *loading* plots. By comparing Figures 2.8 and 2.9, it is clear that high electronegativity corresponds to halogens, while high melting and boiling points correspond to metals.[64]

Score and *loading* plots can answer many questions and may indicate the next step for further multivariate analysis.[63,64]

PCA is not the only tool for Exploratory Data Analysis and Pattern Recognition. Similar tools are Factor Analysis (FA), Cluster Analysis (CA), Linear Discriminant Analysis (LDA) and Soft Independent Modelling of Class Analogy (SIMCA).[63] The target of pattern recognition is *classification*. Applications of pattern recognition are well represented in literature and studies to improve the current methods constitute an active area of the chemometrics research.[40]

2.2.2 Calibration and Multiple Linear Regression

A *regression analysis* is usually performed to study how a measured *independent* variable is related to a *dependent* variable. In this context, the most common application of regression analysis is the construction of calibration lines for analytical instrumentations. For example, in quantitative analysis, it is assumed that a measured independent response/variable (which depends on the analytical technique) is a function of the amount of analyte (dependent variable). The calibration process, via regression, aims to determine and model this function, in order to use this mathematical model to predict the quantities of the same analyte in unknown samples.[8,65]

Calibration has always been a “hot topic” in analytical chemistry and represents an important section of the chemometrics research – many papers on this subject can be found in literature.[40,65] Before the development of chemometrics and multivariate data analysis, the only option was univariate calibration. Univariate calibration is based on only one variable and it is highly sensitive to interferences and matrix changes (the matrix always needs to be the same for calibration standards and samples).[8,66]

The natural alternative to overcome the limitations of univariate calibration, is *multivariate calibration*, where multiple responses/variables are measured simultaneously (i.e. there is not just one independent variable).[8]

The simplest technique for multivariate calibration is MLR.[64] In MLR, one dependent variable is modelled as a function of different independent variables.[63] MLR plays an important role in process control analysis, especially when spectroscopic techniques are in place; e.g. monitoring the absorbance at various wavelengths simultaneously.[65] MLR has also been one of the chemometrics tools successfully implemented to deal with matrix effects and interferences in GFAAS.[60-62]

This technique, although capable of compensating for typical problems encountered in univariate calibration/regression, is the least powerful of the chemometrics calibration/regression techniques. It requires a full knowledge of the *training set* (the composition of all calibration standards must be known), can neglect significant errors and can give problems if there is collinearity between some variables in the model.[63,64]

The techniques discussed in Sections 2.2.3 and 2.2.4 have gained popularity because they can overcome the deficiencies of MLR.

2.2.3 Principal Component Regression

PCR is a combination of PCA and MLR. It is divided into two stages: the first stage consists of a normal PCA to decompose the data matrix (objects/samples and measured variables); the second stage is the regression of the decomposed independent variables. The main advantage of PCR over MLR is that the former can compensate for significant errors.[63,64]

PCR is very useful in spectroscopy of mixtures. If a scientist is interested in the identification of only one compound in a mixture, a PCR can determine this analyte without knowing the other components of the mixture. On the other hand, MLR works properly only when all the significant components are identified.[64]

Collinearity can also be solved with a PCR. First, the regression is performed on *scores* and *loadings* and not on the original variables; second, PCs are orthogonal, therefore collinear variables become independent.[8,65]

The main problem with PCR is that the *scores* used for the regression are not always correlated with the property of interest. For example, if the analyte is at a very low concentration in a sample, in a spectroscopic analysis the main effect will be produced by other constituents that are at higher concentration. These secondary constituents will also be responsible for the largest variation of the analysis and will be represented in the main PCs (such as PC1).[8]

In the next paragraph Partial Least Squares Regression will be discussed, which is considered an improvement of PCR and is becoming the preferred regression technique in chemometrics.[40]

2.2.4 Partial Least Square Regression

PLS-R is the multivariate calibration technique that has been recently dominating the scenario of the chemometrics literature. One of the main reasons for the success of PLS-R is the improvement and availability of commercial software that can apply this technique.[40] PLS-R was also implemented to improve GFAAS performance.[56]

There is much controversy over PLS-R. Some scientists say that PLS-R can be used universally in chemometrics to solve any kind of problem; others say that PLS-R is a useful technique but it cannot be considered universal because most chemometrics problems can be dealt with using PCR.[64]

In PLS-R, the independent variables (usually instrumental responses) are called *predictors* and constitute the *X-block*; the dependent variable (usually concentrations), that must be predicted, is called *predictand* and constitutes the *Y-block*.[8]

PLS-R has all the advantages of PCR. The main difference is that PLS-R assumes that there are errors in both *X*- and *Y-block*, which is very close to reality. For example, if an AAS instrument needs to be calibrated (using the relation between absorbance and concentration), there will be errors during the preparation of the standard solutions and instrumental errors during the reading of the absorbance.[8,64]

PLS-R also includes the data of the *Y-block* in the regression while PCR only takes into account the *X-block*. Therefore, with PLS-R the dependent variable (*Y-block)* has a higher correlation with the property of interest (e.g. analyte concentration).[8,65]

The regression of PLS-R produces PLS components that are similar to the PCs of a PCA (they are often called Principal Components, although this notation is not correct), with a correspondent version of *score* and *loading* plots. However, while in PCA *scores* and *loadings* depend only on the independent variables (*X-block*), in PLS-R they also depend on the dependent variable (*Y-block*).[64]

When there is only one dependent variable in the *Y-block*, as in PCR and MLR, the PLS-R is called PLS-1. An extended PLS version is PLS-2. With PLS-2 different objects are modelled simultaneously (like different analytes at different concentrations).[63,64]

Notwithstanding all its advantages, PLS-R remains a complex technique. In simple situations, such as when all components of a mixture are known, it is generally recommended to apply simpler approaches such as MLR.[64]

2.3 Experimental Design – Where Chemometrics should begin

"From a drop of water a logician could infer the possibility of an Atlantic or a Niagara without having seen or heard of one or the other." [67]
(Sir Arthur Conan Doyle, from "Sherlock Holmes, in a study in scarlet")

The importance of Experimental Design can be symbolised by this quotation by Alan Lakein: "Failing to plan is planning to fail".[68] Accordingly, there are many authors in literature who emphasise the importance of DoE, which has often been neglected by scientists and chemometricians.[8,25,34,63,64,69-74] The message delivered by these authors was already brought up by John Hintermaier in 1948: correct conclusions cannot be obtained from an experiment that has not been properly thought and designed.[75]

The comparison between DoE, as multivariate analysis, and the OVAT (One-Variable-At-a-Time) approach, as univariate analysis, can be associated to the clash between good sense and common sense. If we consider the optimisation of an instrument (or, in general, the study of a system that depends on many variables), we cannot expect to find the best solution by varying one variable, keeping the others at a constant value and ignoring any possible interaction between them. Traditionally, common sense suggests that the OVAT approach is the "logical" way to solve problems; but, especially in a multivariate environment, *common sense is not good sense*.[8,71,76,77]

The concept of Experimental Design was formulated by Sir R. A. Fischer in 1925, almost 50 years before Wold invented chemometrics.[78] From Experimental Design scientists and chemometricians learnt that, before performing any analysis, they must understand the problem and delineate the aim of the experiment. If they neglect this "philosophical" step, they will most probably waste a lot of time in obtaining and analysing useless data that do not contain the sought information ("failing to plan is planning to fail"). In other words, good experimenters know what they want and with a few thought experiments they can obtain the knowledge they seek. Today, this concept is at the basis of the chemometric reasoning.[63,70,71,76]

Experimental Design consists of selecting the best points of the experimental domain by taking into consideration the number of experiments, which is limited by various factors such as time and cost.[78] The main advantages of DoE are summarised below:[63]

- knowing in advance the number of experiments needed to obtain the required information;
- obtaining more information from a smaller number of experiments (in comparison with the OVAT approach);
- obtaining only the needed information;
- investigating more variables simultaneously and studying their combined action (interaction).

Experimental Design is (or should be) an essential chemometric tool that encloses useful methods and techniques. These techniques have been used to solve problems regarding GFAAS, especially interferences and matrix effects.[14,25,26,59-62,72,73,77]

The choice of the "best" DoE technique depends on the objectives and limitations of the experiment and on the type/quantity of information sought – there is no "best design".[63,79] At this regard, there are three main goals in Experimental Design:[8,63,71]

- identify the factors that have the largest effect on the system;
- quantify the effects of these factors and their interactions;
- find the optimum values/conditions for these factors (*optimisation*).

In the following sections, different Experimental Design methods will be illustrated. Before describing the features of these techniques, some basic terminology will be introduced:[8,63,80]

- *System*: it is the process that needs to be studied (e.g. an analytical procedure).
- *Factor*: it is an input variable to investigate. This input may or may not have an influence upon the system. Various factors can be investigated simultaneously. Factors can represent both *quantitative* (temperature,

concentration, etc…) and *qualitative* (catalyst type, solvent nature, etc…) variables.

- *Response*: it is a measure of the influence of a factor on the system (measured output variable). The choice of responses to measure depends on the purpose of the study.
- *Factor Level*: factors can be controlled or uncontrolled. Controlled factors are experimental conditions that can be set at desired values. Each factor can be studied at more than one value and each value is identified as a *level of the factor*.
- *Experimental Domain*: experimental region defined by the levels of the factors. The investigation of the system occurs only within this region.
- *Factor (or Main) Effect*: it is the change of a response caused by the change of the factor levels. It can be both positive (the response increases) and negative (the response decreases). Factors that produce large effects are also defined as important variables.
- *Interaction Effect*: two factors interact when the response of one of them changes according to the level of the other one. An interaction is null when the factors are independent from each other.
- *Order of interaction*: it is determined by the number of factors involved in an interaction.
- *Treatment*: a combination of factors and factor levels.
- *Replicates*: replicated experiments are experiments performed more than once in a design. This means that the whole procedure is replicated, from sample and standard solutions preparation to instrumental analysis. Replication gives more responses for a given treatment and it is used for the estimation of the experimental error.
- *Repeated measurements*: instrumental analyses performed more than once on the same sample. The average value of repeated measurements is normally regarded as a single response. They are performed to determine the instrumental error.

2.3.1 Factorial Designs

A system is a function of the input variables and their interactions. Hence, the first step in an investigation is deciding which factors and which responses will be studied, and one of the challenges is to assess how (or if) different variables influence each other.[71]

Factorial designs are a renowned class of DoE well represented in literature.[25,73,77,81,82] The word factorial, in this case, does not have the mathematical meaning of an integer multiplied by all integers smaller than itself (e.g. 3! = 3×2×1); but it means that a system is studied by varying different factors simultaneously, to reveal their effects and interactions. Thus, all possible combinations of different factors and levels (treatments) can be investigated with a number of experiments that depends on the number of factors and their levels. For example, if one has to investigate three factors **A**, **B** and **C**, with **A** at 2 levels, **B** at 3 levels and **C** at 4 levels, the total number of treatments would be 2×3×4 = 24. Factors are represented by capital letters, while interactions are indicated by the association of the correspondent letters: **AB** indicates the second-order interaction between factors **A** and **B** and so it is for **AC** and **BC**; **ABC** represents the higher, third-order interaction between **A**, **B** and **C** (higher-order interaction are usually negligible).[8,70]

By means of an analysis of variance (ANOVA) each main effect and interaction can be quantified. If replicates are included, also the experimental error can be estimated. From this, an *empirical model* can be obtained, and via this model it is possible to estimate an unknown response within the experimental region.[8]

Generally, not all factors can be controlled (set at the desired level) by the experimenters, such as room temperature and humidity, origin of certain samples or analyst on duty. Uncontrolled conditions can produce effects that can be confounded with the factor effects. There are two main approaches to compensate for these uncontrolled variables: *Randomisation* and *Blocking*.

Randomisation is a basic experimental principle. It prevents that unknown and undesirable factors get confused with the main effects. It simply consists of

performing the experimental runs in a *random* order. It is essential when the experimental conditions can vary in time during the course of the investigation, such as room temperature and humidity on different days or different times of the day.[63,71,76]

Blocking differs significantly from randomisation. Randomisation is used to eliminate the effect of factors that *might* influence the responses; that is, it is unknown if, how and when these factors will have an effect on the experiments. On the other side, when there is an external factor that *will* affect the system, but one is not interested in quantifying it by including it in the design, its effect can be removed by arranging the treatments in *blocks*.[71] A block is a homogeneous set of experiments. Each block is characterised by uniform conditions as it "contains an equal amount" of the disturbing factor, while all the design factors are equally represented in each block. Typical situations where blocking is required are:[63,70,76]

- when experiments must be carried on different days;
- when there are different operators or different instruments;
- when raw material from different batches/suppliers is used.

2.3.2 2-Level (2^k) Factorial Designs

To investigate the effect of a factor, one needs to change its level and observe how the response is affected by this change; this means that a factor needs at least 2 levels to be studied. The simplest and most common form of DoE is represented by the 2-Level Factorial Designs (2^k-FD), where all factors are investigated at 2 levels, *Low* and *High* level, which are coded as **+** and **-** (or **+1** and **-1**). Coding is very important as it provides a *dimensionless* coordinate system.[8,70,71,76]

Levels can also be assigned to qualitative (or *discontinuous*) factors. For quantitative (or *continuous*) factors the Low and High levels are numbers. For qualitative variables the levels represent "choices" – e.g. type of catalyst/reagent; presence or absence of catalyst/reagent; sample preparation method; internal standard element; etc…

For a k number of factors, the number of treatments needed for a *complete* 2^k-Factorial Design (CFD) is $\mathbf{2^k}$. With 3 factors, for example, the experimental domain would be studied with 8 experiments (2^3).[8,70,71]

A 2^k-FD is characterised by a *design matrix*, $\mathbf{2^k \times k}$, where the experiments are listed. The columns denote the factors and the rows (*runs*) the experiments. The signs **+** and **-** show the levels of each factor in a particular treatment. Table 2.2 is an example of design matrix for a CFD with 3 factors. For example, in Experiment 1 all factors are at the high level (+) and in Experiment 6 factors **A** and **C** are at their low level (-) and factor **B** is at its high level (+). Likewise, a CFD with 2 factors would produce a design matrix with 2 columns and 4 runs, a design with 4 factors would produce a matrix with 4 columns and 16 runs, and so forth.[8] Note that the runs are arranged according to the *standard* order, but an experimenter should perform the experiments randomly (see Paragraph 2.3.1).

Table 2.2: Example of a Design Matrix for a 2 levels-3 factors CFD.

Run	Factor A	Factor B	Factor C
1	+	+	+
2	-	+	+
3	+	-	+
4	-	-	+
5	+	+	-
6	-	+	-
7	+	-	-
8	-	-	-

There is a general rule to construct a 2^k design matrix. The first column alternates 1 plus and 1 minus sign, the second column alternates 2 plus and 2 minus signs, the third column alternates 4 plus and 4 minus signs, the fourth column alternates 8 plus and 8 minus signs, and so forth. When the experiments have been performed and the responses recorded, it is common practice to add another column for listing the response obtained from each treatment.[8,71]

A 2^k-FD can also be represented geometrically, where the geometrical figure symbolises the experimental domain. The easiest examples are those with 2 and

3 factors, which are represented by a square and a cube, respectively. The human mind cannot visualise the geometrical figure formed by designs with more than 3 factors, but more cubes can be combined to represent designs with 4 factors or more.[63,70]

Figure 2.10 shows the geometrical representations of a 2^k-FD with 2 and 3 factors (Figures 2.10a and 2.10b, respectively). The arrows symbolise the change from low to high level. Each vertex (sphere) represents a treatment and the factor levels of that particular experiment are shown in the square brackets. This representation shows how the number of factors and their levels define the experimental domain. Additionally, replacing each sphere with the value of the correspondent response helps in identifying that factor/level combination that has the highest effect.

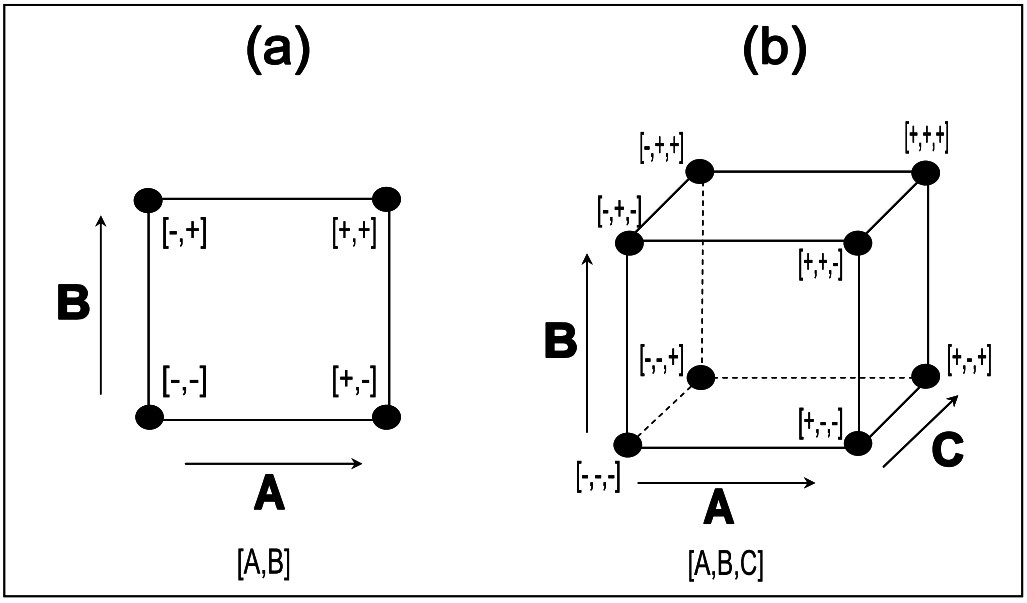

Figure 2.10: Geometrical representation of a 2-Level Factorial Design with (a) 2 factors (A and B) and (b) 3 factors (A, B and C). The square brackets show the factor levels of each treatment (each vertex is a treatment).

2^k-FD can estimate the main and interaction effects. A statistical evaluation of these effects can be done via an ANOVA. However, to perform an accurate ANOVA, an estimation of the experimental error is needed and replicated experiments have to be performed.[8,71] The addition of a replicated *centre*

point (with all factors at their middle level, placed at the centre of the square/cube) can also be used to estimate the experimental error.[63,76,83] However, an experimenter is often not interested in a statistical evaluation and only needs to know which factors and interactions have the most prominent effects. This can be achieved without the addition of replicated or centre experiments, by treating the magnitude of the highest-order interaction as the experimental error. Furthermore, replicated experiments and centre points increase the total number of experiments and time and cost of the investigation.[8,70]

The *magnitude* of an effect is a relative value that comes from the coding (+/-) of the factor levels. It represents the extent of the change that it is observed in a response when a factor level is moved from Low to High. The magnitude of factor and interaction effects can be easily illustrated by means of graphical tools, such as *Pareto Charts* and *Normal Probability Plots*.[8,71]

A **Pareto chart** shows factor and interaction effects ordered by magnitude. In the example below (Figure 2.11) the vertical axis lists all factors and interactions, while the horizontal axis indicates their magnitude.[8]

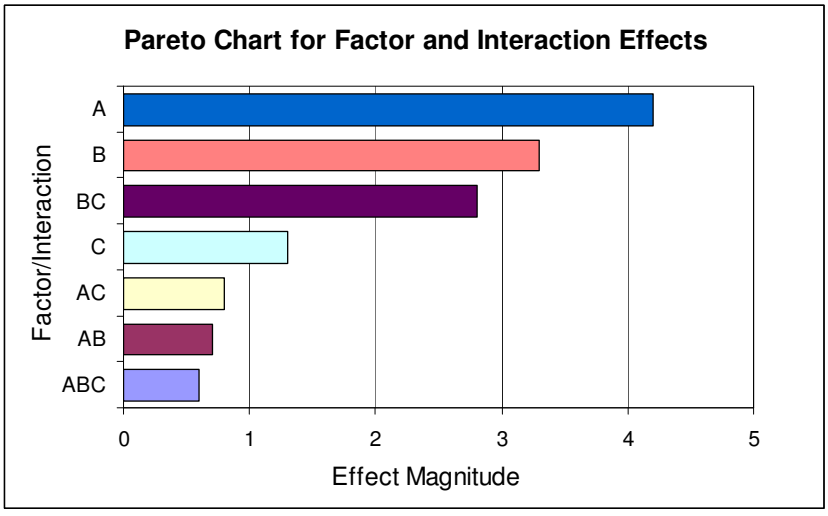

Figure 2.11: Example of Pareto chart. The vertical axis lists the factors and interactions ordered by magnitude; the horizontal axis shows the magnitudes.

In Figure 2.11, factors **A** and **B** and the interaction **BC** have the strongest effects. The interactions **AC** and **AB** have a magnitude very close to the interaction **ABC**. If **ABC** is the experimental error, **AC** and **AB** are solely background noise, hence meaningless. The magnitude of the effect of Factor **C** is higher than **ABC**, but it is probably not significant if compared with **A**, **B** and **BC**. An ANOVA could confirm if Factor **C** is significant or not by means of the P-Value and F-Test: if $P < 0.05$ and $F > F_{value}$ the effect is considered significant and must be taken into account.[63,84]

Table 2.3 is a fictitious and simplified ANOVA table that shows only the F-Ratios and the P-Values. This fabricated table was designed to explain the Pareto chart in Figure 2.11 and it shows the correlation between the factors with the highest magnitude in the Pareto chart and their P-value in the ANOVA table.

Table 2.3: Fictitious and simplified ANOVA table showing only F-Ratio and P-Value.

Source of Variation (Factor/Interaction)	F-Ratio	P-Value
A	1197.43	0.000
B	906.05	0.001
C	9.32	0.073
AB	2.67	0.120
AC	1.84	0.202
BC	852.19	0.008
ABC	1.07	0.305

A **Normal Probability Plot** (N-Plot) is another graphical illustration used to identify important factors/interactions. N-plots are derived from the normal distribution. The *x*-axis represents the value (or magnitude) of the effects; the *y*-axis represents the "normality" of the points (percentiles). If the points (effects) are all normally distributed they will follow a straight line. Factors/interactions that produce significant effects will not be normally distributed and will be distant from the straight line (like outliers). The farther they are from the central region of the straight line, the more important they are.[63,71]

An example of N-Plot is shown in Figure 2.12 (adapted from Trindade *et al.*, 2006).[82] In this example, there are two points (red circles) that do not belong to

the straight line formed by the other points and therefore represent significant factors. If one should associate this N-Plot with the examples in Figure 2.11 and Table 2.3 the red circles would represent factors A and B.

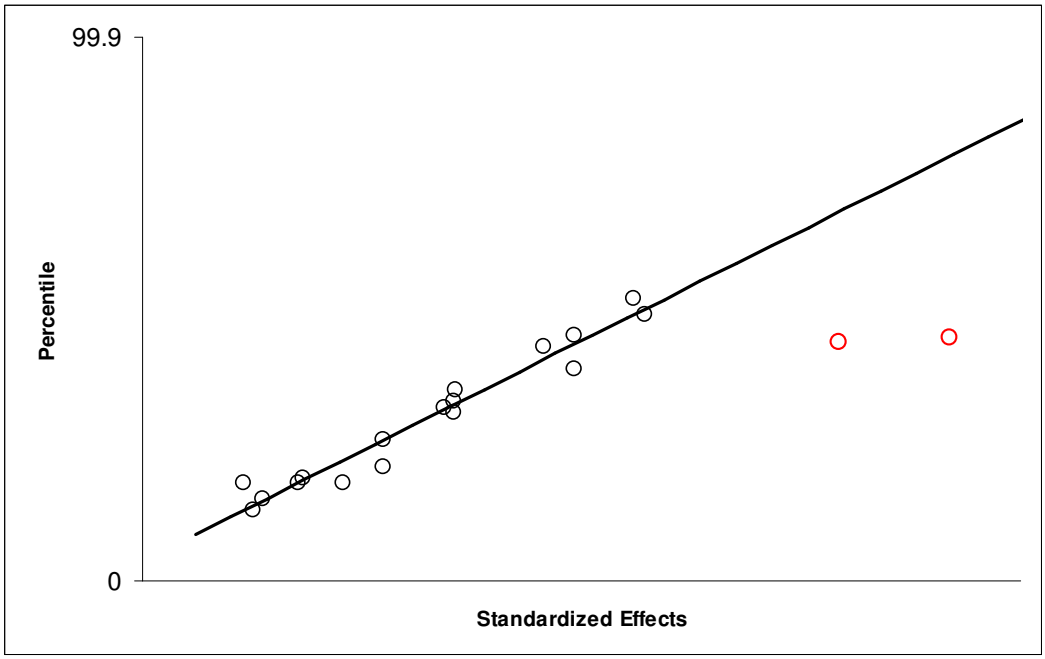

Figure 2.12: Example of a Normal Probability Plot (adapted from Trindade et al., 2006).[82]

2.3.3 Screening

CFD are practical only when the number of factors is limited, as the number of experiments increases exponentially with the number of variables (2^k). When a large quantity of factors needs to be studied, a *screening* analysis is performed to identify the important ones. Following their identification, the important variables can be studied with a more detailed design. This approach leads to time and cost savings.[76,83,84]

There are different screening techniques. The most popular approaches are Fractional Factorial Designs (FFD), described in Section 2.3.4, and Plackett-Burman (PB) designs, discussed in Section 2.3.5.

2.3.4 Fractional Factorial Designs

As it was mentioned above, with a CFD, the number of treatments increases dramatically with the number of factors (e.g.: $2^5 = 32$, $2^7 = 128$ and $2^{10} = 1024$ experiments). However, in these situations, it is generally not necessary to perform every treatment as it is possible to obtain the desired information by performing just a *fraction* of the experiments and sacrificing the high-order interaction effects (in certain cases, the design can be completed at a later stage by adding more experiments). First, the magnitude of high-order interaction effects is typically very small and it gets smaller when the order increases; hence, it does not make sense to perform many experiments just to quantify these meaningless effects. Second, when there are many factors, one is usually not interested in the high-order interactions as the important variables have not been identified yet.[8,70,71]

Factorial designs that use this short cut are known as FFD and are defined by $\mathbf{2^{k-p}}$ experiments. They are the economical alternative to CFD and are also known as *reduced* factorial designs because they reduce the number of experiments. They can be considered as blocks taken from a CFD.[8,76,84]

Confounding is the price to pay for performing fewer experiments and it is the basic principle of FFD. With confounding, interactions are *confounded* (or *aliased*) with the main effects. This means that these interactions cannot be studied independently as their effects are included in the magnitude of the main effects. This is generally not a problem, because the magnitude of the main effects is much larger and the contribution from the interactions can be considered negligible.[63]

For example, a 2^4 CFD can estimate all main effects (**A**, **B**, **C**, and **D**) and interactions (**AB**, **AC**, **AD**, **BC**, **BD**, **CD**, **ABC**, **ABD**, **ACD**, **BCD** and **ABCD**) with 16 experiments. A fractional design 2^{4-1}, which involves only 8 experiments, can estimate all main effects, which are confounded with the negligible high-order interactions (**A**+**BCD**, **B**+**ACD**, **C**+**ABD** and **D**+**ABC**). On the other hand, low-order interactions become aliases of one another (**AB**+**CD**, **AC**+**BD**, **AD**+**BC**) and their effects cannot be distinguished. For example, if the effect **AB**+**CD** is

important, one cannot say if it is due to **AB**, **CD**, or both, unless more experiments are performed. 2^{4-1} is defined as the *half fraction* of 2^4 because it requires only half of the experiments (all $\mathbf{2^{k-1}}$ FFD are half fractions of $\mathbf{2^k}$). The aliased effects constitute the *confounding pattern* (a detailed illustration on how confounding patterns are formed is beyond the scope if this dissertation and it can be found in literature).[70,71]

Table 2.4: Example of a 2^{4-1} Fractional Factorial Design.

Run	Factor A	Factor B	Factor C	Factor D
1	+	+	+	+
2	-	+	+	-
3	+	-	+	-
4	-	-	+	+
5	+	+	-	-
6	-	+	-	+
7	+	-	-	+
8	-	-	-	-

Table 2.5: Lists of complete and fractional factorial designs (2^{k-p}) with resolution (CFD indicates a non-fractionated design) and number of experiments (runs).

k	p	Design	Resolution	Runs
3	0	2^3	CFD	8
3	1	2^{3-1}	III	4
4	0	2^4	CFD	16
4	1	2^{4-1}	IV	8
5	0	2^5	CFD	32
5	1	2^{5-1}	V	16
5	2	2^{5-2}	III	8
6	0	2^6	CFD	64
6	1	2^{6-1}	VI	32
6	2	2^{6-2}	IV	16
6	3	2^{6-3}	III	8
7	0	2^7	CFD	128
7	1	2^{7-1}	VII	64
7	2	2^{7-2}	IV	32
7	3	2^{7-3}	IV	16
7	4	2^{7-4}	III	8
8	2	2^{8-2}	V	64
9	4	2^{9-4}	IV	32
10	6	2^{10-6}	III	16

Table 2.4 is an example of a 2^{4-1} FFD design matrix. The first 3 columns are constructed as for a 2^3 design. The fourth column is derived by multiplying (algebraically) the signs of the first 3 columns. This explains why **D** and **ABC** are aliased: they are calculated in the same way (**D** = **ABC = A×B×C**).[70,71] Table 2.5 lists a series of complete and fractional factorial designs, with their resolution (see below) and number of experiments.

Confounding can have different patterns, especially when the number of factors is large. The degree of confounding defines the *resolution* of the fractional design. The most common resolutions are III, IV and V.[71]

Resolution III Designs cannot distinguish between main and second-order interactions effects. They can only estimate main effects and assume that all interactions are negligible. These designs are the most economical ones and are used mainly for screening. Resolution III designs are *saturated* designs, as they require the minimum number of experiments for a given number of factors. Examples of resolution III designs are 2^{5-2}, 2^{6-3}, 2^{7-4} and 2^{10-6} FFD.[63,71,76]

Resolution IV Designs can separate main from second-order interaction effects. Main effects are aliased with the effects of third-order interactions, which are usually negligible, but second-order interactions cannot be separated from each other. These designs can quantify main effects and give an indication of which second-order interaction is significant. Examples of resolution IV designs are 2^{4-1}, 2^{6-2}, 2^{10-5} and 2^{10-4} FFD.[63,71,76]

Resolution V Designs can estimate and quantify main and second-order interaction effects. The main effects are confounded with fourth-order interaction effects, while second-order interactions are aliased with third-order interactions. These designs offer almost the same accuracy of a CFD. Examples of resolution V designs are 2^{5-1}, 2^{8-2} and 2^{10-3} FFD. For a given number of factors, the resolution decreases when p increases (2^{k-p}).[63,71,76]

2.3.5 Plackett-Burman Saturated Designs

PB designs where created by Robin Plackett and Peter Burman to develop new weapons during World War II – an example of how the War stimulated the minds of scientists. PB designs are a particular class of saturated 2-level designs and a very powerful tool for preliminary screenings and robustness tests. These designs can study the effects of $k+1$ factors with only k experiments, but only when $k+1$ is a multiple of 4. For example, PB designs can study 3 factors with 4 experiments, 7 factors with 8 experiments, 11 factors with 12 experiments, and so forth; FFD require always 2, 4, 8, 16, 32, etc., experiments, regardless the number of variables (Table 2.5).[8,70,78] The "economy" of these designs is the reason for their popularity.[85-88]

PB designs are characterised by very complex confounding patterns as they assume that all types of interactions can be ignored. They are very useful to identify unimportant factors. Nonetheless, they should be used mostly to decrease the initial number of factors as it is not recommended to use PB designs to choose the final important variables. Another characteristic of these designs is that the number of factors can also be smaller than $k+1$ with the remaining factors (needed to get a multiple of 4) set as "dummy" variables; "dummy" variables can then be used to estimate the experimental error.[63,71,84]

PB designs are also characterised by design matrices. Although the mechanisms to construct these matrices is quite different from the 2^k-FD matrices (for further details, please refer to the references).[71,84]

2.3.6 Optimisation Designs

The term *optimisation* has always been particularly significant in chemistry (especially analytical chemistry). Typically, the object of an optimisation is to find the experimental conditions that will lead to the best outcome. Historically, this task has been performed following the OVAT approach which is time consuming and unable to estimate interactions.[89]

The main outcome of CFD or screening designs is the identification of the factors that have a significant effect on the system. In general, this is not enough to find the combination of levels that maximises the response. An *Optimisation Design* investigates the key factors (preferably 2 to 5 variables) at more than 2 levels. This way, a clearer idea of the system can be obtained together with the *optimal* factor levels that lead to the *optimum* response. In certain cases, optimisation designs can be considered as extended 2^k-FD. They are necessary when strong interactions and/or non-linear relationships are expected, as with 2^k-FD it is not possible to detect curvatures.[63,70,76,83,85-87]

Since optimisation designs employ more than 2 levels, they cannot be used to study qualitative factors, as a choice or a quality cannot be set at intermediate levels – think of the mid-level between Analyst A and Analyst B or between the presence and the absence of a catalyst. These qualitative variables needs to be fixed at their "best level" according to the results obtained from the screening.[63]

The following section introduces two popular optimisation designs: Central Composite Design (CCD) and Box-Behnken Design (BBD).

2.3.7 Central Composite and Box-Behnken Designs

Central Composite Designs are the combination of a factorial design (2^k) with a different type of design, *viz.* a *star* design ($2k$+1), with a total number of $\mathbf{2^k + 2k + 1}$ runs. To avoid a large number of experiments, these designs should be employed only with 5 variables or less. CCD can estimate main and interaction effects. They can also estimate the error when the centre point is replicated. As factorial designs are based on 2 levels and star designs on 3 levels, CCD have 5 levels. The high star levels are higher than the factorial points while the low star levels are lower than the factorial ones.[63,89] A 3-Level Factorial Design (3^k-FD) can also detect curvatures, although the number of experiments required would be higher than with a CCD (Table 2.6).[70]

Table 2.6: Comparison of number of experiments between 3^k Factorial and Central Composite Designs.

Number of factors	Number of experiments	
	3^k FD	CCD
2	9	9
3	27	15
4	81	25
5	243	43

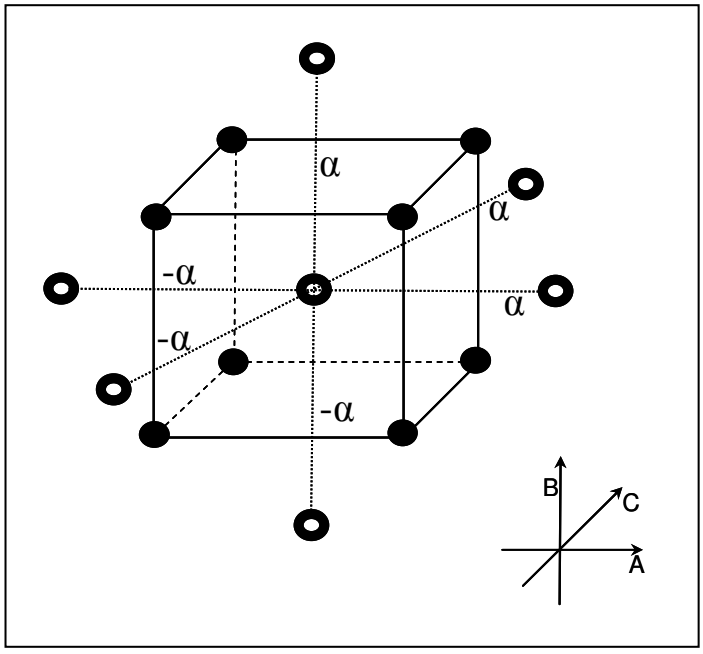

Figure 2.13: Geometrical representation of a Central Composite Design with 3 factors (A, B and C). The solid spheres represent the factorial design and the hollow spheres the star design points; α and -α indicate the distance of the star points from the centre (adapted from Esbensen, 2002).[63]

The differences between 2^k-FD (CFD and FFD) and CCD can be appreciated by examining the geometrical representation (Figure 2.13) and the design matrix (Table 2.7) of a CCD with 3 factors. In Figure 2.13 the solid spheres belong to the factorial design (at the vertices of the cube) and the hollow spheres to the star design, where α is the distance of the star points from the centre of the cube. This picture highlights the combination of the factorial design with the star design and illustrates how a factorial design can be extended to a CCD by adding

the star points.[63,71] In Table 2.7, runs 1 to 8 define the factorial design, while runs 9 to 15 define the star design. The coded values for the star design are defined by α. In this design the centre point (Run 9: 0, 0, 0) is not replicated.

Table 2.7: Design Matrix of a generic Central Composite Design

Run	Factor A	Factor B	Factor C
1	+1	+1	+1
2	-1	+1	+1
3	+1	-1	+1
4	-1	-1	+1
5	+1	+1	-1
6	-1	+1	-1
7	+1	-1	-1
8	-1	-1	-1
9	0	0	0
10	α	0	0
11	$-\alpha$	0	0
12	0	α	0
13	0	$-\alpha$	0
14	0	0	α
15	0	0	$-\alpha$

The distance of the star points from the centre, α, defines the type of CCD. Two special types of CCD are formed when $\boldsymbol{\alpha = 1}$ or $\boldsymbol{\alpha = \sqrt{k}}$.

When $\boldsymbol{\alpha = \sqrt{k}}$, all points (star and factorial) have the same distance from the centre and are located on the surface of a sphere (Figure 2.14a). This situation is characterised by *rotatability*, which implies that the experimental error depends only on the distance from the centre and the importance of each treatment is the same.[63,71]

When $\boldsymbol{\alpha = +/-1}$, the star points are located in the centre of each cube face (Figure 2.14b) and the CCD is called Face-Centred Cube (FCC) design. FCC designs are very useful when there are constraints on the choice of the levels, such as when it is not possible to go higher and lower than the high and low factorial levels. One advantage of this situation is that only 3 levels are required and the number of experiments is reduced.[63,71]

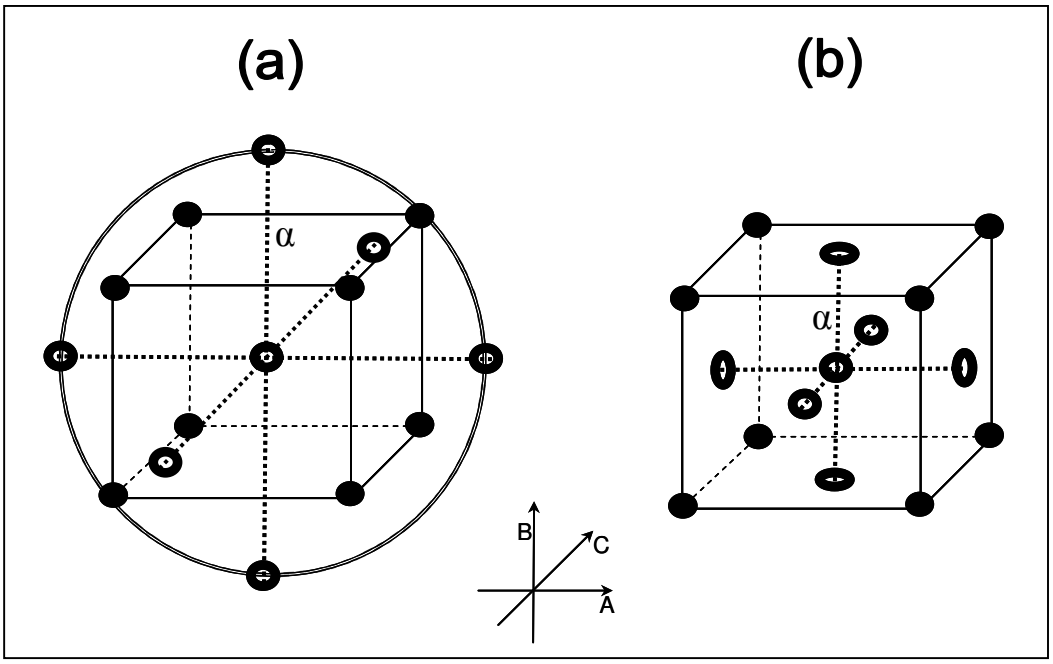

Figure 2.14: Example of special CCD with 3 factors (A, B and C): a) Rotatable CCD where all points have the same distance from the centre (α = √3) and are placed onto the surface of a sphere that circumscribes the cube; b) Face Centred Cube (α = 1) with the star points placed at the centre of each cube face (adapted from Esbensen, 2002).[63]

CCD can also be formed by combining a star design with an FFD. In this way the number of experiments decreases but the complexity of the results increases.[71]

Box-Behnken Designs (BBD) are the economical alternative to CCD. These designs analyse each factor (not less than 3) at 3 levels. By using only 3 levels, instead of 5, the number of experiments decreases dramatically. Figure 2.15 shows the geometrical representation of a BBD. From this figure two main observations can be made: (i) the points are distributed spherically around the centre of the cube and (ii) no point is placed at the vertices of the cube. The first observation implies that also this design is rotatable. From the second observation other key characteristics can be deduced:

- BBD do not include factorial points and cannot be used to extend and optimise 2^k-FD;

- BBD have no level combination where all factors are at their low or high level (no treatment at the vertices), thus extreme experimental conditions are avoided;
- as with FCC designs, no point is set outside the cube.

The results obtainable from a BBD, at a lower cost, are not always less reliable than the results obtainable from a CCD.[63,71]

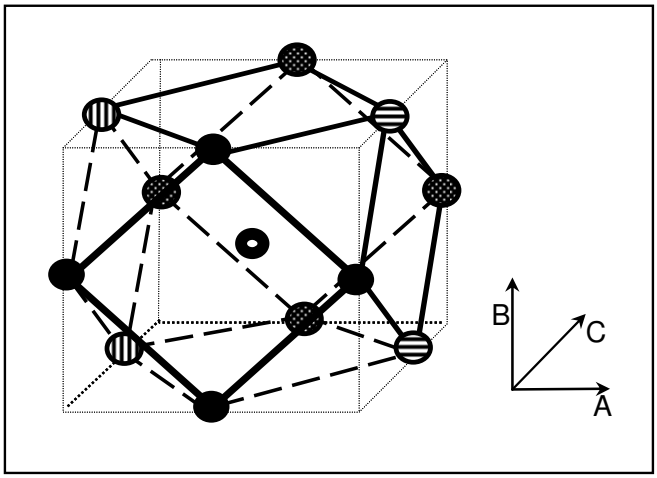

Figure 2.15: Geometrical representation of a Box-Behnken Design with 3 factors (A, B and C). The different shadings aid to visualise the position of the spheres in the cube. The hollow sphere represents the centre experiment at the centre of the cube (adapted from Esbensen, 2002).[63]

Optimisation designs are generally associated with **Response Surface plots** (from Response Surface Methodology, RSM), another graphical tool for the evaluation of the results (Section 2.3.2). A response surface plot is a 3-dimensional graph produced by plotting 2 interacting variables and the response. This graph shows which level combination yields the best output. These plots need a curvature estimation that cannot be obtained from 2^k-FD. Response surfaces can also be seen from above, in 2-dimensional graphs, as **contour plots** (composed of *isoresponse* curves).[63]

Figure 2.16 is a 3-dimensional response surface plot. In this example, the highest response is obtained when factors A and B are close to their middle level.

Figure 2.17 shows the correspondent contour plot, which indicates the region of maximum response at the centre of the plot.

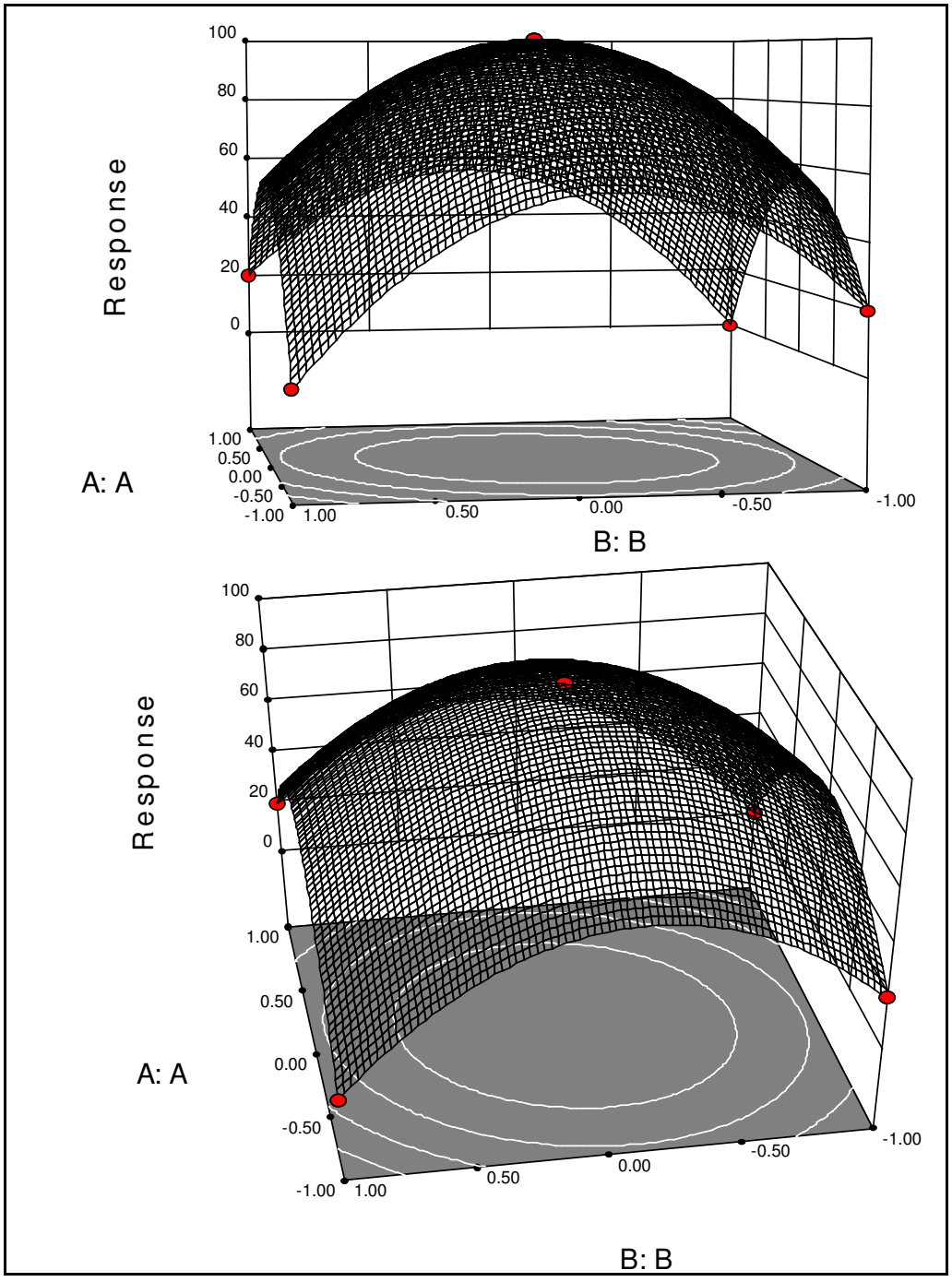

Figure 2.16: Example of response surface plot. The highest response is obtained when both factors (A and B) are set around their middle level.

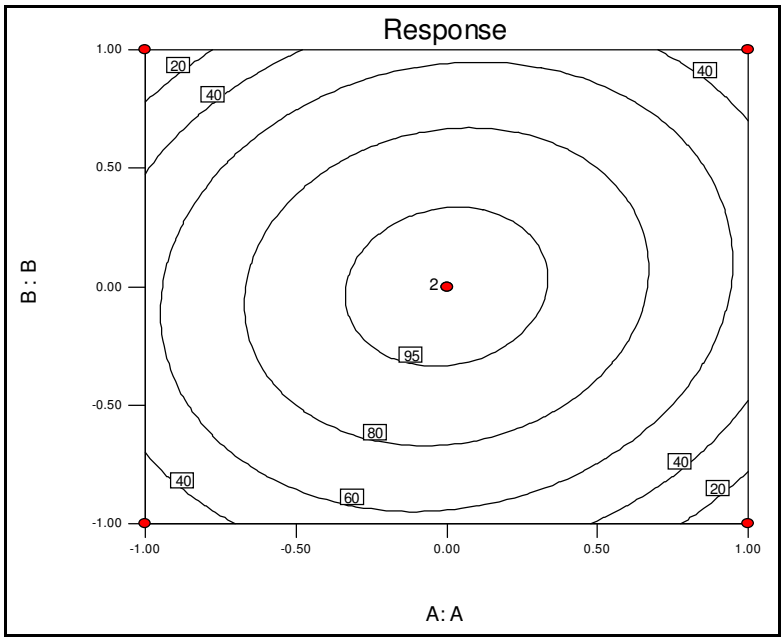

Figure 2.17: Example of contour plot. The isoresponse curves enclose a region of maximum response.

Other optimisation techniques are available for more specific problems and applications. The most popular methods are mentioned below (for a detailed description of these methods, the reader is referred to the specialised literature).[8,71,89]

Simplex Approach (or Design): it is a sequential design adopted when the "distance" between the optimum region and the experimental domain in unknown. It consists of performing a series of experiments aimed at ruling out the experimental conditions (level combinations) that do not yield the desired results. The main disadvantages of this method are the large number of experiments and the impossibility of introducing replicates.

Doehlert Designs: in this design the experimental points are allocated on the surface of a sphere (instead of a cube). Each factor can be studied at different levels and the total number of experiments is given by $\boldsymbol{k^2 + k + 1}$. This design is very useful when not all factors can be set at the same number of levels or when some factor needs to be studied in more detail than others.[90,91]

Mixture Designs: these designs are implemented when the proportions of the factors are restricted or fixed – e.g. the sum of the ingredients of a cup cake cannot be higher than 100g (or 100%).

3 EXPERIMENTAL SECTION

"No amount of experimentation can ever prove me right; a single experiment can prove me wrong."[92]
(Albert Einstein)

3.1 Instruments and accessories

The experiments were performed using a *GBC® Avanta UltraZ* (Ver1.7) GFAAS instrument, with Zeeman background correction. Ridged tube furnaces were used to prevent the sample from spreading out of the ends. The GFAAS instrumental parameters are listed in Table 3.1. The lamp current is the current at which the Hollow Cathode Lamp operates and determines the intensity of the lamp; 343.5 nm is one of the wavelengths at which Rh absorbs lights; 0.2 nm is a typical setting for the slit width to obtain good sensitivity and minimise noise and interferences; the background correction is based on Zeeman Effect and a magnetic field of 1.1 T is generally strong enough to correct for interferences and non-specific absorption.

Table 3.1: Instrumental parameters

Lamp current (mA)	Wavelength (nm)	Slit width (nm)	Background correction	Magnetic Field Strength (T)
6	343.5	0.2	On	1.1

Triplicate instrumental measurements and the average of these readings was treated as the experimental response. The sampling and instrumental operations are described in Section 2.1.2.

A Millipore® water deioniser system, with Milli-Q® and Elix® 3 components, was used to generate high purity deionised water.

High magnitude pictures of used graphite tubes (Section 4.5) were obtained using a Zeiss® Ultra 55 field emission gun scanning electron microscope (FEGSEM).

The experiments (Sections 4.1 to 4.4) were designed using the software Design-Expert® Version 7.1.6 (Copyright © 2008). The analysis of the results was performed with Design-Expert® Version 8.0.1 (Copyright © 2010).

Statistica® Version 8.0 (Copyright© Statsoft, Inc. 1984-2007) was employed to perform a PCA of the results (Section 4.6).

3.2 Reagents

The samples for the experiments were prepared using the following reagents:

- GR Toluene for analysis (Merck);
- GR Methanol for analysis (Merck);
- Octanal for synthesis (Merck);
- n-Hexane *chromasolv* (Sigma-Aldrich);
- ROPAC powder (CAS No: 25470-96-6) was used to prepare the Rh samples in the organic solvents above;
- High purity deionised water (from Millipore® system);
- ICP standard containing 1000 mg/l Rh in HNO_3 2-3% (Merck) was used to prepare the Rh sample in water.

3.3 Sample preparation

All samples, with the exception of the water sample, were prepared by dissolving 0.239 g of ROPAC in approximately 30 mL of warm toluene. This mixture was transferred into a 100 mL flask which was filled to the mark with toluene, giving a stock solution of 0.5 mg/mL of rhodium. From this stock solution, 1 mL was diluted to 100 mL with toluene, giving a solution containing 0.005 mg/mL of rhodium. Finally, 0.5 mL of this intermediate stock solution was pipetted into a 50 mL flask and diluted with one of the organic solvents listed above, forming a sample containing 50 ng/mL (ppb) of Rh.

The water sample was prepared starting from an ICP standard containing 1000 mg/l Rh in HNO_3 2-3% (Merck). From this stock solution, 1 mL was diluted in 100 mL of deionised water forming an intermediated stock solution of 10 mg/l (ppm) of Rh. From this intermediate solution, 0.5 mL were diluted to 100 mL to give a 50 ppb solution of Rh in water.

After the sample preparation, five solutions of 50 ppb of Rh were obtained in hexane, toluene, methanol, octanal and water.

3.4 Experimental procedure

The experiments performed in this study followed a stepwise procedure, where each step depended on the results of the previous step. Hence, the details of the experimental designs and the investigated parameters are discussed together with the results, in Chapter 4.

4 RESULTS AND DISCUSSION

"Facts are the air of scientists. Without them you can never fly."[93]
(Linus Pauling)

Figures 4.1 and 4.2 show the matrix effects that were observed during the GFAAS method development for Rh analysis. In GFAAS, the ashing step is supposed to remove the matrix of samples; hence, a preliminary test was performed to investigate the ashing parameters and assess the relationship between the ashing process and matrix effects (Section 4.1).

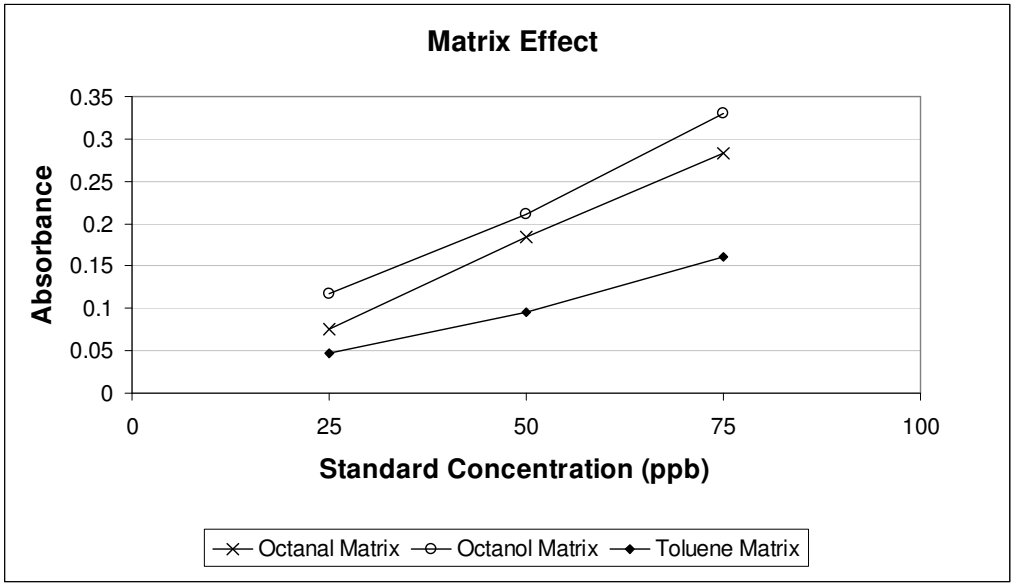

Figure 4.1: Difference in sensitivity/absorbance between octanal, octanol and toluene due to matrix effects in GFAAS.

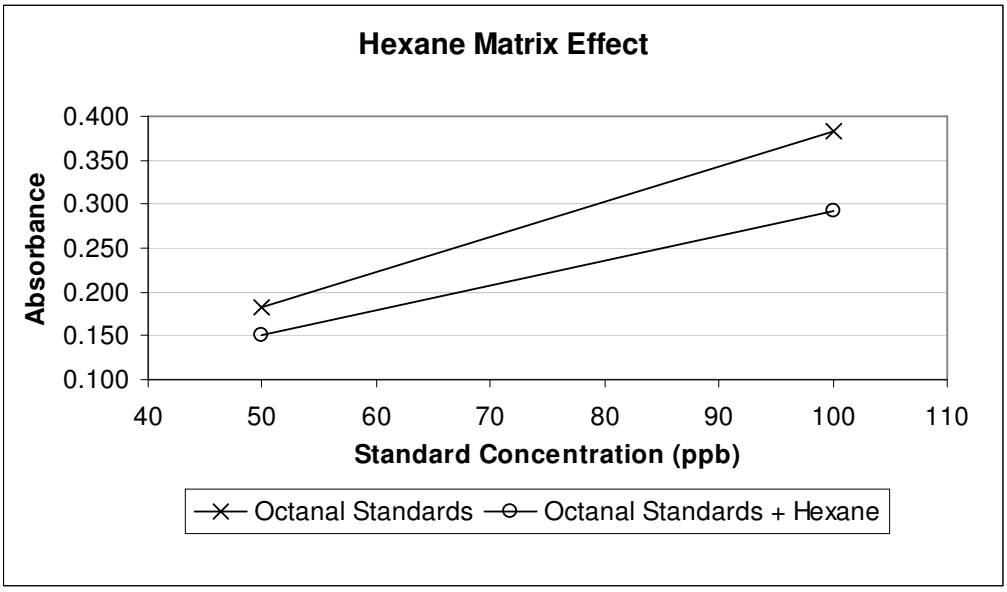

Figure 4.2: Hexane matrix effect: the sensitivity of standard solutions prepared in a mixture of octanal and hexane is lower than the sensitivity of standard solutions prepared in pure octanal.

4.1 Preliminary test

Based on the actual furnace programme implemented for the industrial Rh analysis (Table 4.1), a 2^3 CFD (with the addition of a centre point) was proposed to uncover the relationship between matrix effects and ashing step parameters, hold time (Ht_{Ash}), ramp time (Rt_{Ash}) and final temperature (T_{Ash}) – the design matrix is shown in Table 4.2. The levels of the experimental domain were arbitrarily chosen to enclose the ashing settings of the furnace programme in Table 4.1.

Table 4.1: Furnace programme implemented for actual Rh analysis.

Step	*Final Temp. (ºC)*	*Ramp Time (s)*	*Hold Time (s)*	*Inert gas*
Pre-heating	40	1.0	2.0	On
Injection				On
Drying	70	30	40	On
Transition	200	20	20	On
Ashing	1000	40	10	On
Stabilisation	1000	0.1	2.0	Off
Atomisation (signal reading)	2400	1.0	1.5	Off
Cleaning	2650	1.0	1.5	On

Table 4.2: 2^3 CFD for the investigation of the ashing parameters.

Run	Levels	T_{Ash} (ºC)	Rt_{Ash} (s)	Ht_{Ash} (s)
1	+++	1100	40	15
2	-++	700	40	15
3	+-+	1100	15	15
4	--+	700	15	15
5	++-	1100	40	5
6	-+-	700	40	5
7	+--	1100	15	5
8	---	700	15	5
9	000	900	27.5	10

Table 4.3: Preliminary test experiments and results. Abs1 and Abs2 are the averages of triplicate absorbance readings.

Run	Levels	T_{Ash} (ºC)	Rt_{Ash} (s)	Ht_{Ash} (s)	Abs 1	Abs 2
1	+++	1100	40	15	0.601	0.615
2	---	700	15	5	0.613	0.604
3	000	900	27.5	10	0.601	0.612

GFAAS is a slow technique and each analysis/experiment requires a significant amount of time. Therefore, to obtain a rapid estimation of this relationship, only the experiments at the margins of the experimental domain were performed (Table 4.3). These experiments consist of the treatments where all factors are set at their high level (+++) or low level (---). Furthermore, the treatment with the factors at their mid level (000) was also chosen. Hence, only 3 experiments from the design (composed of 9 treatments) were performed. This can also be illustrated with the geometrical representation of the design (Figure 4.3) where the chosen treatments are highlighted by red circles.

This preliminary test consisted of analysing a Rh sample by changing the ashing settings according to Table 4.3 (the settings for the other programme steps were not modified). The test was performed twice and the results are shown in Table 4.3 (the results are the averages of triplicate measurements). By looking at these results, one can easily see that the differences, in terms of GFAAS, are insignificant. This means that there is no direct or strong correlation between the sensitivity of the instrument and the ashing parameters (at least within these

levels) and that probably the matrix effects do not depend on the ashing step (at least not on this step alone). Hence, a broader DoE was proposed to investigate all the relevant furnace factors (Section 4.2).

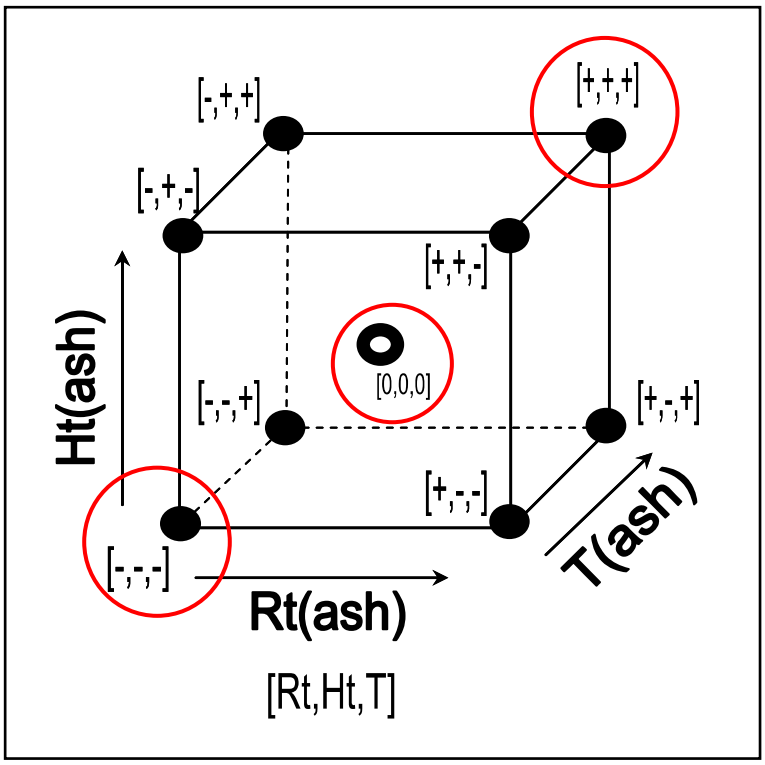

Figure 4.3: Geometrical representation of the 2^3 CFD proposed to study the ashing parameters. The hollow sphere indicates the centre point and the red circles show the experiments chosen for the preliminary test.

4.2 First screening of GFAAS parameters – 2^{7-4} FFD

Not all furnace parameters can be included in an experimental design (Table 4.1). The atomisation temperature is an intrinsic feature of each element; therefore, changing this parameter will definitely affect the sensitivity, which might be misinterpreted as matrix effects. The atomisation hold time should be kept at a minimum to preserve the lifetime of the graphite tubes and provide good sensitivity by producing narrow peaks. The drying step is optimised by visual inspection through a micro-camera that shows the inside of the tube during the analysis; the drying parameters are set (optimised) when a complete and smooth evaporation is observed for all solvents. Hence, seven furnace factors were identified:

1 Transition step final temperature ($\mathbf{T_{Tr}}$), between 80 and 200ºC;
2 Transition step ramp time ($\mathbf{Rt_{Tr}}$), between 5 and 20 seconds;
3 Transition step hold time ($\mathbf{Ht_{Tr}}$), between 5 and 20 seconds;
4 Ashing final temperature ($\mathbf{T_{Ash}}$), between 300 and 1000ºC (the high and low levels were reduced after the results from the preliminary test were obtained);
5 Ashing ramp time ($\mathbf{Rt_{Ash}}$), between 5 and 40 seconds (the low level was reduced after the results from the preliminary test were obtained);
6 Ashing hold time ($\mathbf{Ht_{Ash}}$), between 2 and 10 seconds (the high and low levels were reduced after the results from the preliminary test were obtained);
7 Atomisation ramp time ($\mathbf{Rt_{Atom}}$), between 0.8 and 4 seconds.

A complete factorial design with 7 factors (a 2^7 CFD) would require 128 experiments (Section 2.3.2). However, there is no need to perform a complete factorial design to screen the available factors and identify the important ones (Section 2.3.3). Fractional factorial designs are a cost effective way to study 7 factors or more (Section 2.3.4). With a 2^{7-4} FFD only 8 experiments are needed. The proposed 2^{7-4} FFD to screen the 7 furnace parameters is shown in Table 4.4.

Four samples containing 50 ppb of Rh were prepared in Hexane, Methanol, Octanal and Toluene (Section 3.3) – at this stage of the project, the Millipore system was not available and it was not possible to include the Water sample. The 4 samples were subjected to the 2^{7-4} FFD, resulting in a total of 32 experiments.

GFAAS is not a very robust technique. The same sample, analysed at different times, can produce different raw results. Hence, with GFAAS it is necessary to perform a calibration before every analysis and check the calibration regularly. In light of this, to obtain a statistically significant dataset, it was necessary to perform all experiments in the shortest period of time (as the results were raw absorbance readings, not subjected to/corrected by a calibration). Octanal and Hexane were analysed on the same day, while Toluene and Methanol were

analysed three days later. The results obtained from these experiments are shown in Table 4.5 (the results are the averages of triplicate absorbance readings).

Table 4.4: 2^{7-4} FFD for the investigation of the relevant furnace parameters.

Treatment	T_{Tr} (ºC)	Rt_{Tr} (s)	Ht_{Tr} (s)	T_{Ash} (ºC)	Rt_{Ash} (s)	Ht_{Ash} (s)	Rt_{Atom} (s)
1	80	5	5	1000	40	10	0.8
2	80	20	5	1000	5	2	4
3	80	5	20	300	5	10	4
4	80	20	20	300	40	2	0.8
5	200	5	5	300	40	2	4
6	200	20	5	300	5	10	0.8
7	200	5	20	1000	5	2	0.8
8	200	20	20	1000	40	10	4

Table 4.5: Results from the 2^{7-4} FFD for Hexane, Methanol, Octanal and Toluene. The results are the averages of triplicate absorbance readings.

Treatment	Abs			
	Methanol	Octanal	Toluene	Hexane
1	0.261	0.171	0.174	0.063
2	0.194	0.115	0.139	0.027
3	0.193	0.149	0.147	0.032
4	0.239	0.195	0.210	0.081
5	0.214	0.151	0.129	0.025
6	0.225	0.187	0.188	0.027
7	0.255	0.184	0.183	0.048
8	0.216	0.154	0.155	0.014
Average	0.225	0.163	0.166	0.040

N-Plots and Pareto charts (Section 2.3.2) were generated from the analysis of the results of the 2^{7-4} factorial designs. These graphs are shown in Figures 4.4 to Figure 4.12. Each factor is identified with a capital letter: A-Rt_{Tr}; B- Ht_{Tr}; C-T_{Tr}; D-Rt_{Ash}; E-Ht_{Ash}; F-T_{Ash}; G-Rt_{Atom}. In this discussion, the *x*-axis of the Pareto charts lists the different factors (*rank*) while the *y*-axis indicates the magnitude (*normal % probability*). In the Pareto charts, very important factors have a magnitude higher than the red line (which indicates the Bonferroni limit), while non important factors have a magnitude lower than the black line (the t-Value limit).[94] Magnitudes between the red and black lines need to be evaluated with the aid of the N-Plots and ANOVA tables.

These graphical results show that for Methanol, Octanal and Toluene there is only one significant factor, i.e. the atomisation ramp time. Although the N-Plots of the Methanol (Figure 4.4) and Octanal (Figure 4.6) samples might suggest the presence of another important factor – i.e. ashing ramp time (D) for Methanol and ashing temperature (F) for Octanal – this is invalidated by the correspondent Pareto charts (Figures 4.5 and 4.7, respectively) that show that only the Rt_{Atom} factor (G) has a significant magnitude. Figures 4.8 and 4.9 show the high importance of the atomisation ramp time for the Toluene sample. On the other hand, no significant factor was identified for the Hexane sample (Figures 4.10 and 4.11).

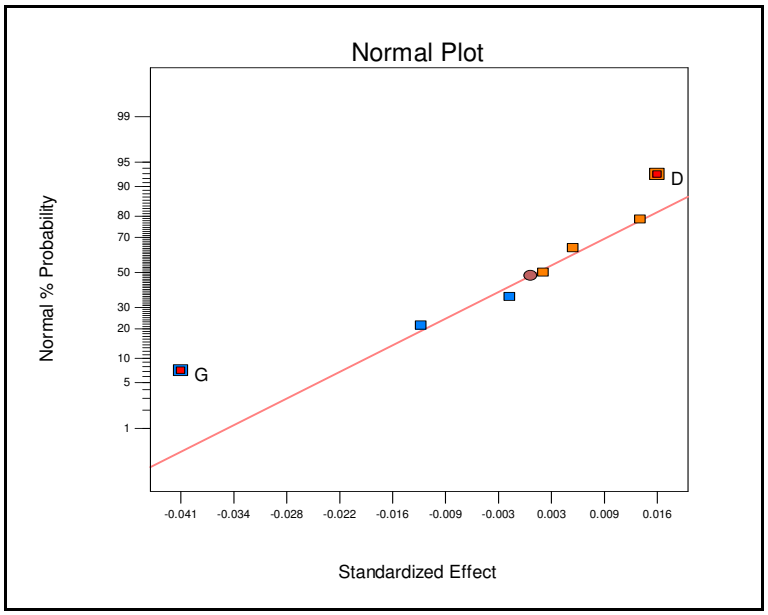

Figure 4.4: N-Plot obtained from the 2^{7-4} FFD applied to the Methanol sample. The two significant factors (outside the straight line) are Rt_{Atom} (G) and Rt_{Ash} (D).

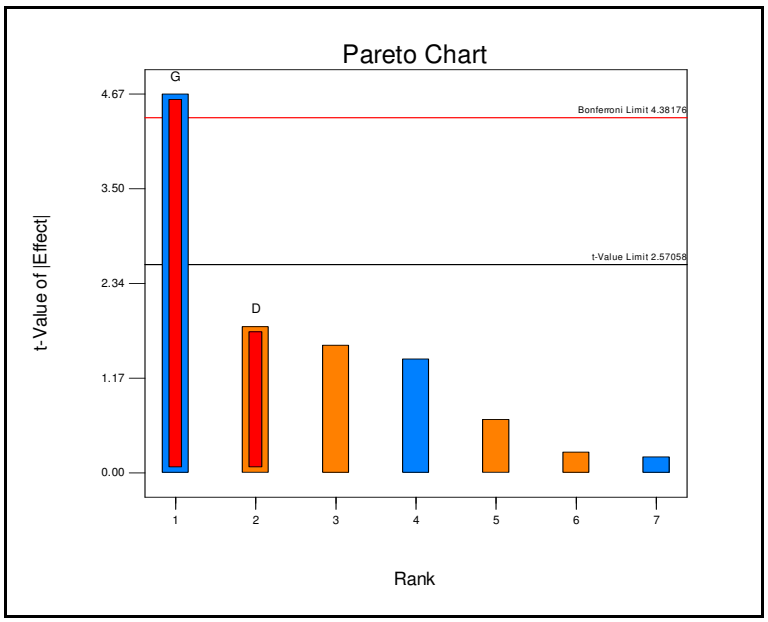

Figure 4.5: Pareto Chart obtained from the 2^{7-4} FFD applied to the Methanol sample. The only significant factor is Rt_{Atom} (G).

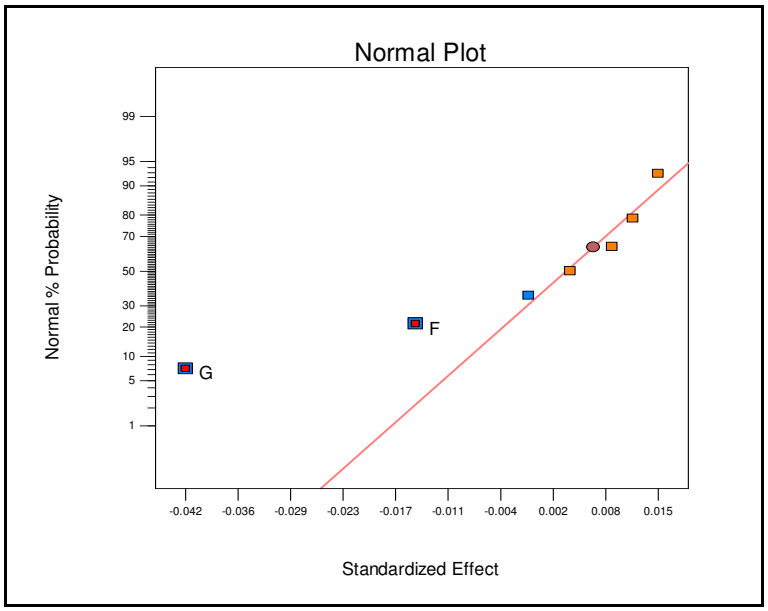

Figure 4.6: N-Plot obtained from the 2^{7-4} FFD applied to the Octanal sample. The two significant factors (outside the straight line) are Rt_{Atom} (G) and T_{Ash} (F).

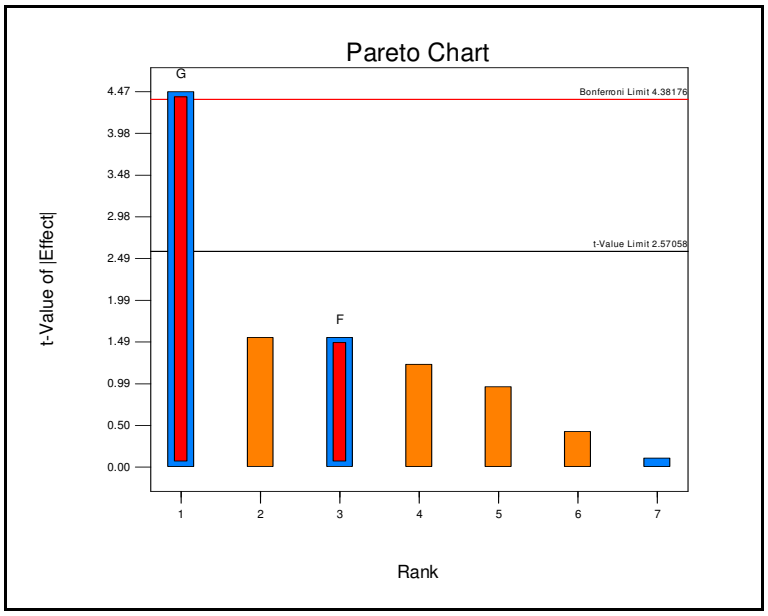

Figure 4.7: Pareto Chart obtained from the 2^{7-4} FFD applied to the Octanal sample. The only significant factor is Rt_{Atom} (G).

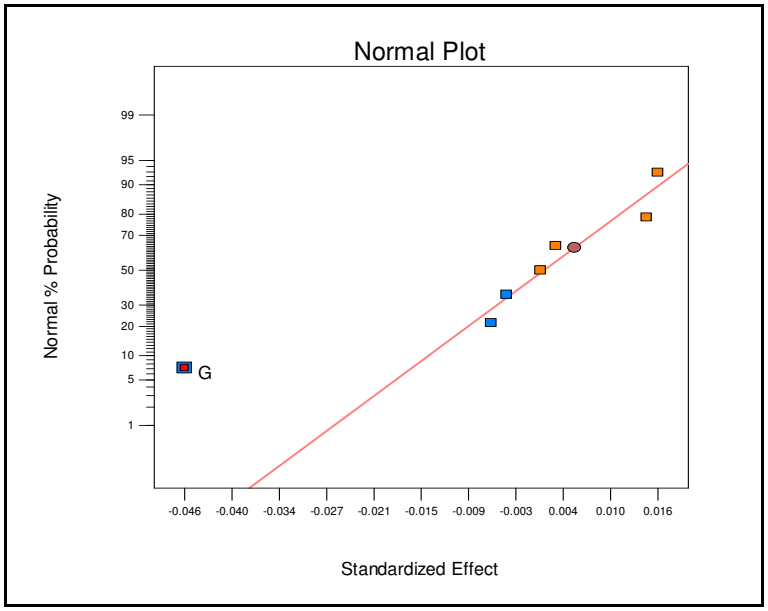

Figure 4.8: N-Plot obtained from the 2^{7-4} FFD applied to the Toluene sample. The only significant factor (outside the straight line) is Rt_{Atom} (G).

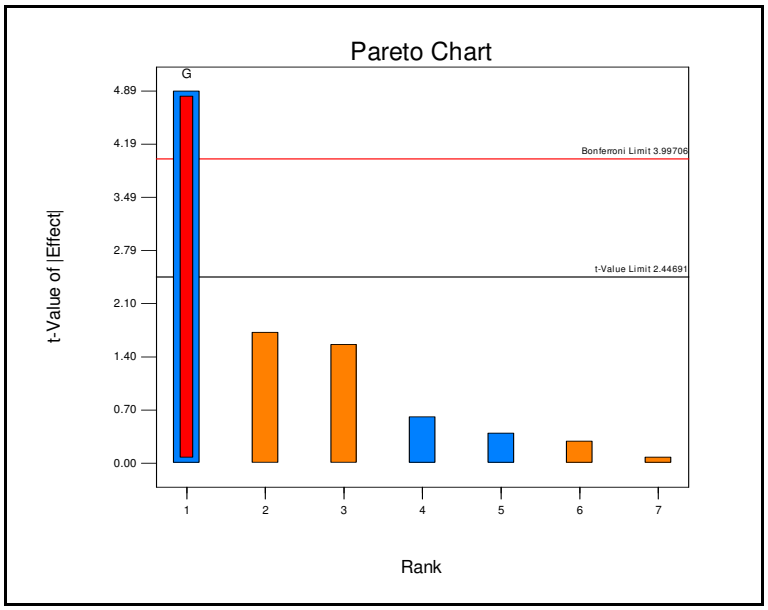

Figure 4.9: Pareto Chart obtained from the 2^{7-4} FFD applied to the Toluene sample. The only significant factor is Rt_{Atom} (G).

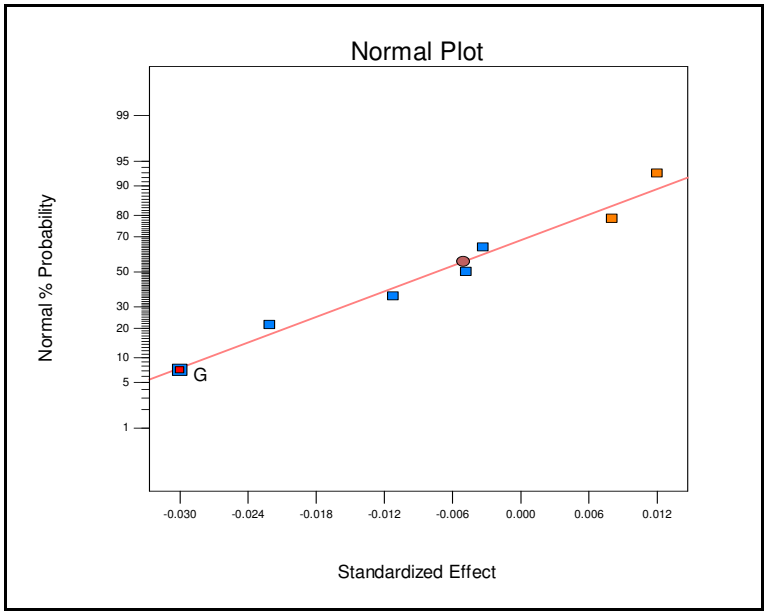

Figure 4.10: N-Plot obtained from the 2^{7-4} FFD applied to the Hexane sample. No significant factor is displayed.

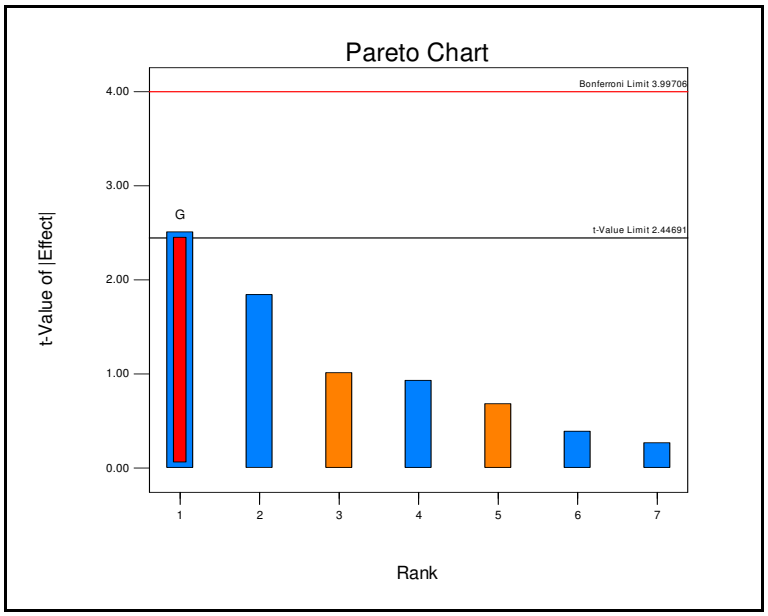

Figure 4.11: Pareto Chart obtained from the 2^{7-4} FFD applied to the Hexane sample. No factor has a magnitude significantly higher than the others.

The high importance of the atomisation ramp time should not be considered an unexpected result. The atomisation ramp time can significantly affect the atomisation process; hence, it can have a large impact on the sensitivity of the method. This is confirmed by the shape of the absorbance peak, which changes dramatically when the atomisation ramp time is set at 4 seconds (Figure 4.12b). When an analysis is performed under optimal conditions, it typically produces a peak that resembles the one portrayed in Figure 4.12a. A signal such as the one in Figure 4.12b is usually associated with low sensitivity. This kind of signals is observed when the atomisation is not fast enough and the atoms are removed through diffusion faster than they are formed by atomisation.

Following the reasoning above, the treatments of the factorial design where the Rt_{Atom} factor is set at 4 seconds could be associated with inaccurate results and may not be statistically sound. To perform an accurate analysis, this factor should be set at a level that would allow the formation of peaks such as the one in Figure 12a. The deletion of this factor prior the data analysis is not a suitable solution because the factor would simply be replaced by one of its aliased terms (which would show the highest significance; Section 2.3.4 and

Appendix B). In addition, due to the non-robust nature of GFAAS, it is not possible to repeat only certain treatments to correct the design, as those experiments might experience a signal/sensitivity drift. For the same reason, it is not recommended to apply the same design to the Water sample and compare the results with the other samples that were analysed considerably earlier.

For example, in Table 4.5, there are no apparent matrix effects between the Octanal and the Toluene sample as their results are practically the same. This is only due to the fact that the two sets of experiments were performed on different days. However, the scope of this project was not to assess the most sensitive solvent.

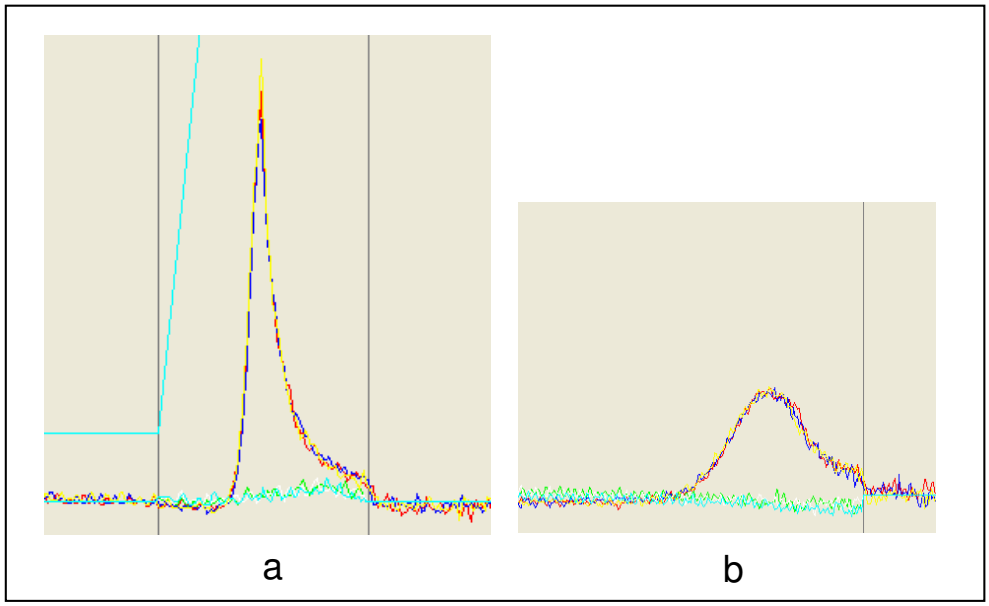

Figure 4.12: a) Peak shape produced with optimal GFAAS conditions (e.g. with 0.8-1.0 seconds atomisation ramp time); b) peak shape observed when the atomisation ramp time is 4.0 seconds. The three different colours (red, yellow and blue) represent triplicate measurements.

A design without the Rt_{Atom} factor (a 2^{6-3} FFD) was proposed. This design was also applied to the Water sample. The results obtained from this experimental design are described in Section 4.3.

4.3 Second screening of GFAAS parameters – 2^{6-3} FFD

The design matrix of the 2^{6-3} FFD is shown in Table 4.6. This experimental plan was applied to 5 samples – Methanol, Octanal, Toluene, Hexane and Water – resulting in a total of 40 experiments. The results obtained from these experiments are listed in Table 4.7 (the results are the averages of triplicate absorbance readings).

In Table 4.7 the significant difference between the results obtained from the Octanal and Toluene samples can be seen, confirming the presence of matrix effects (compare Tables 4.5 and 4.7). In this case, these sets of experiments were performed on the same day.

Table 4.6: 2^{6-3} FFD for the follow up investigation of the relevant furnace parameters.

Treatment	T_{Tr} (ºC)	Rt_{Tr} (s)	Ht_{Tr} (s)	T_{Ash} (ºC)	Rt_{Ash} (s)	Ht_{Ash} (s)
1	80	5	5	1000	40	10
2	80	20	5	1000	5	2
3	80	5	20	300	5	10
4	80	20	20	300	40	2
5	200	5	5	300	40	2
6	200	20	5	300	5	10
7	200	5	20	1000	5	2
8	200	20	20	1000	40	10

Table 4.7: Results from the 2^{6-3} FFD for Hexane, Methanol, Octanal, Toluene and Water. The results are the averages of triplicate absorbance readings.

Treatment	Abs				
	Methanol	Octanal	Toluene	Hexane	Water
1	0.197	0.210	0.143	0.049	0.182
2	0.198	0.187	0.149	0.009	0.103
3	0.213	0.220	0.164	0.022	0.197
4	0.203	0.224	0.162	0.027	0.212
5	0.210	0.227	0.162	0.025	0.191
6	0.213	0.269	0.162	0.028	0.189
7	0.198	0.208	0.161	0.011	0.184
8	0.204	0.204	0.128	0.025	0.174
Average	0.205	0.219	0.154	0.025	0.179

The N Plots and Pareto charts generated from these experimental designs are shown in Figures 4.13 to Figure 4.22.

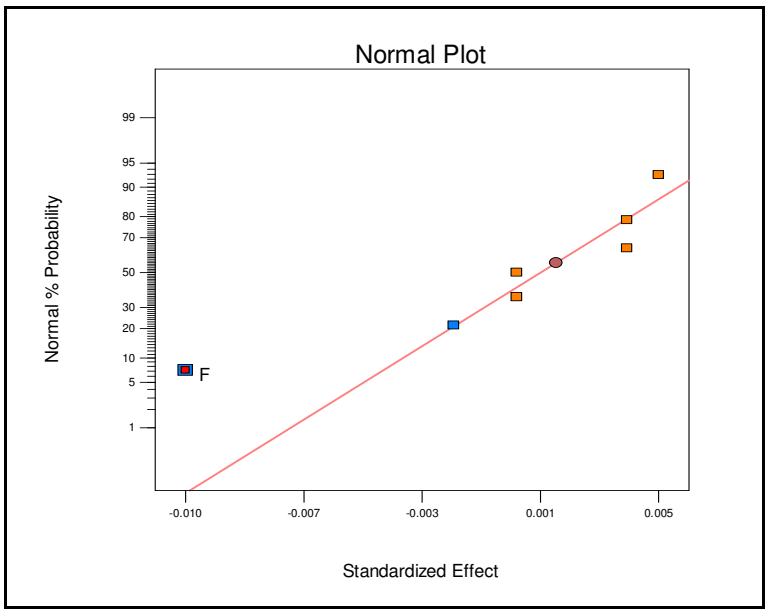

Figure 4.13: N-Plot obtained from the 2^{6-3} FFD applied to the Methanol sample. The only significant factor is T_{Ash} (F).

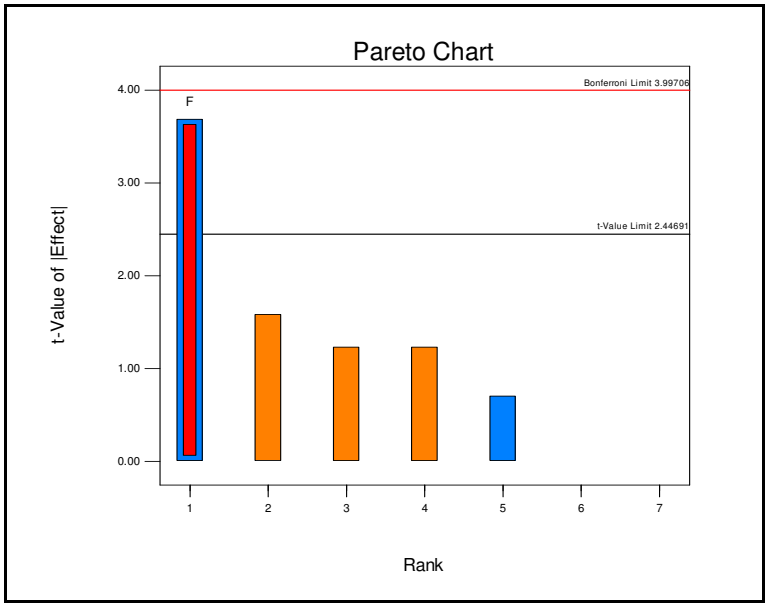

Figure 4.14: Pareto Chart obtained from the 2^{6-3} FFD applied to the Methanol sample. The only significant factor is T_{Ash} (F).

The reader should note that the factorial designs discussed in this Chapter did not include any replicate or centre point (with the exception of the DoE on the atomisation step in Section 4.4); consequently, no real experimental error was produced.

The Methanol sample (Figures 4.13 and 4.14) represents the only case where the omission of the Rt_{Atom} factor uncovered the importance of a different parameter, i.e. the ashing final temperature (F). Although the magnitude of this factor is not above the Bonferroni limit (Figure 4.14), its significance is confirmed by the P-value (i.e. 0.0103 < 0.05) in the correspondent ANOVA table (Table A.5 in Appendix A). This result could confirm that the matrix effects are controlled and/or generated during the ashing step.

The same parameter also appears to be important for Octanal and Toluene (Pareto charts in Figures 4.16 and 4.18, respectively). However, its magnitude is not high enough to be considered statistically significant. Furthermore, the correspondent N-Plots (Figure 4.15 for Octanal and Figure 4.17 for Toluene) show that no significant factor was highlighted by the experimental designs as all points follow a straight line; although, the P-values in the ANOVA tables (Table A.6 for Octanal and Table A.7 for Toluene) are still within the limits of significance (0.0403 and 0.0460, respectively; less than 0.05). For these reasons (no factor is significantly more important than the others), it is not possible to obtain a mathematical model that could correct for the matrix effects when Octanal and Toluene are employed.

A similar situation is observed for the Hexane sample. The N-Plot (Figure 4.19) highlights two distinct factors: ashing ramp time (D) and ashing hold time (E). This result could be used to correct the matrix effects when Methanol and Hexane based samples are analysed, by correcting T_{Ash} for Methanol samples and Rt_{Ash} and Ht_{Ash} for Hexane samples. However, by examining the Pareto chart, one can see that these two factors are below the t-Value limit and, consequently, they are not significant. This is also verified from the ANOVA table (Table A.8) that shows a P-value of 0.0557 for Rt_{Ash} and 0.0693 for Ht_{Ash} (higher than 0.05). This leads to the conclusion that for hexane based

samples it is also not possible to obtain a mathematical model that could correct for matrix effects.

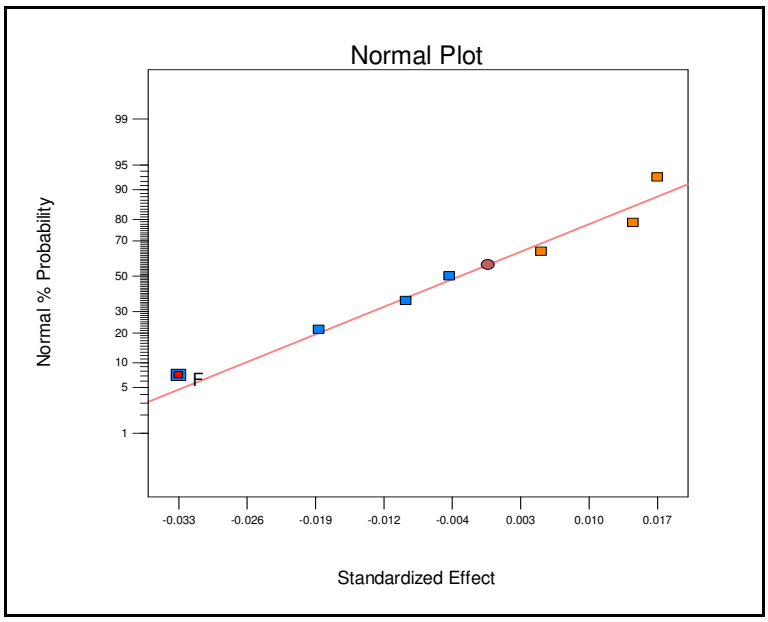

Figure 4.15: N-Plot obtained from the 2^{6-3} FFD applied to the Octanal sample. No significant factor is displayed.

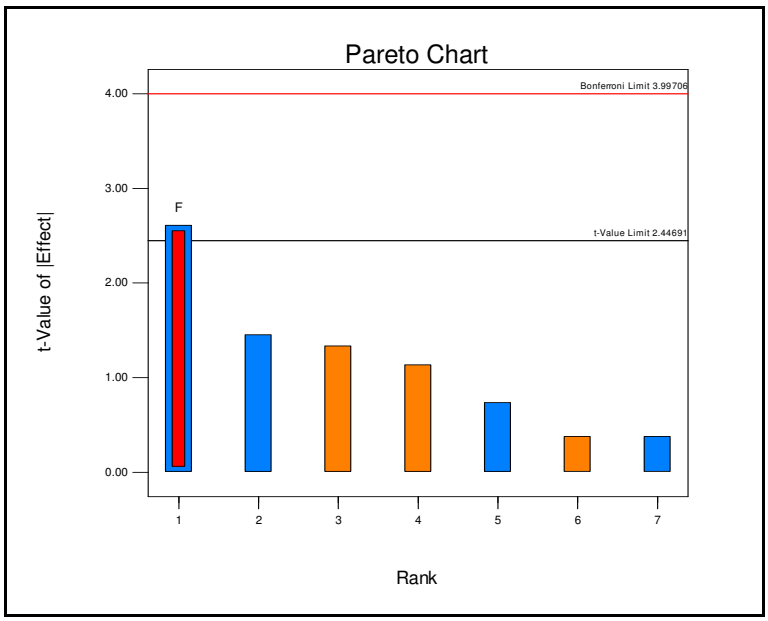

Figure 4.16: Pareto Chart obtained from the 2^{6-3} FFD applied to the Octanal sample. T_{Ash} (F) has a higher magnitude than the other factors although it should not be considered a significant factor (compare with Figure 4.15).

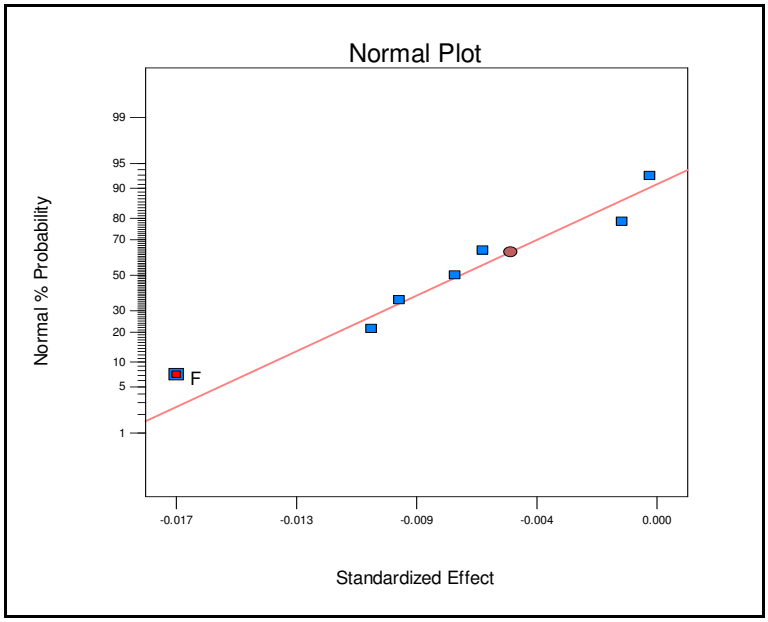

Figure 4.17: N-Plot obtained from the 2^{6-3} FFD applied to the Toluene sample. No significant factor is displayed.

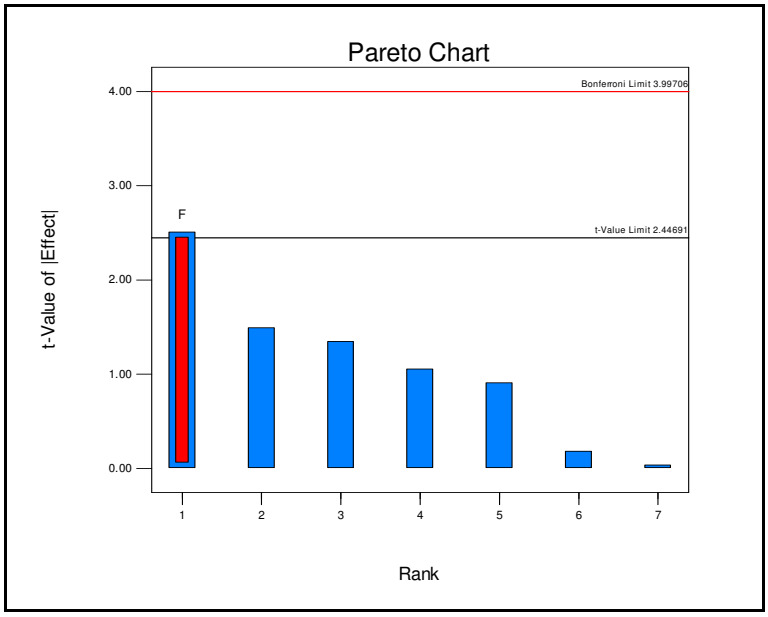

Figure 4.18: Pareto Chart obtained from the 2^{6-3} FFD applied to the Toluene sample. T_{Ash} (F) has a higher magnitude (rank) than the other factors although it cannot be considered a significant factor (compare with Figure 4.17).

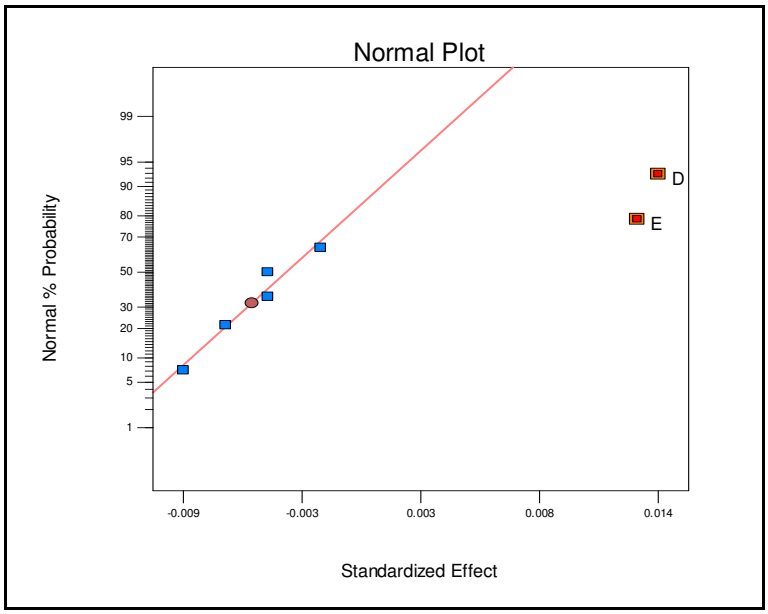

Figure 4.19: N-Plot obtained from the 2^{6-3} FFD applied to the Hexane sample. The factors outside the straight line are Rt_{Ash} (D) and Ht_{Ash} (E); nevertheless, they are not significant factors when the correspondent Pareto chart is considered (Figure 4.20).

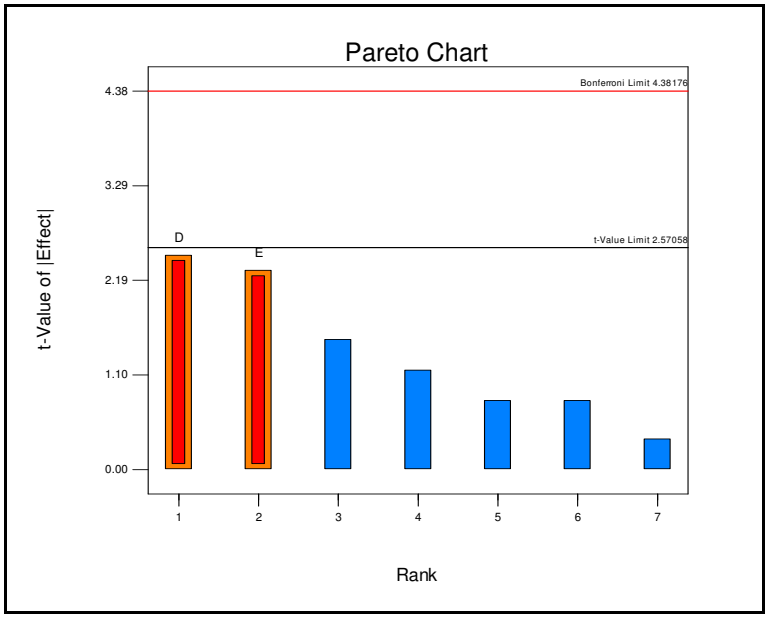

Figure 4.20: Pareto Chart obtained from the 2^{6-3} FFD applied to the Hexane sample. No factor displays a significantly high magnitude.

The N-Plot and the Pareto chart obtained from the Water sample (Figures 4.21 and 4.22) clearly show the absence of any important parameter. This is also in accordance with the P-value: 0.1187 (Table A.9). Similarly, as for Octanal, Toluene and Hexane, a mathematical model cannot be obtained for Water.

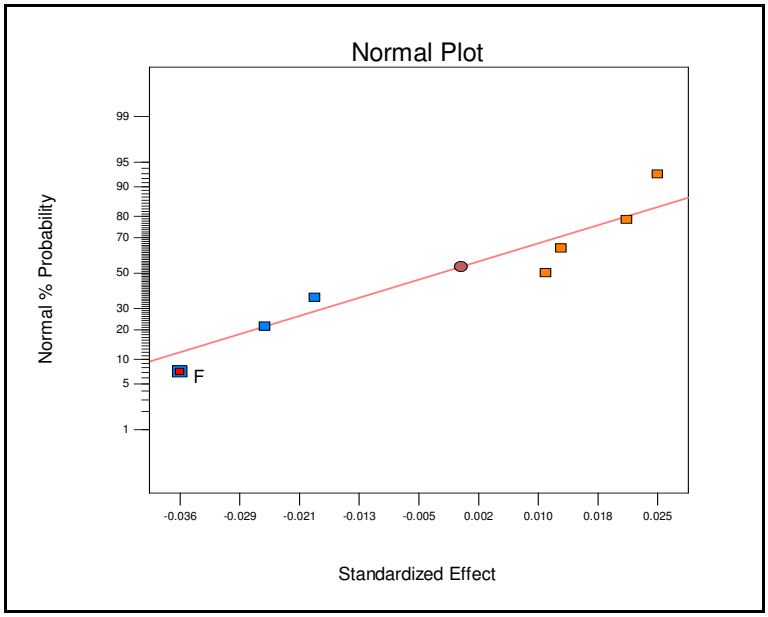

Figure 4.21: N-Plot obtained from the 2^{6-3} FFD applied to the Water sample. No significant factor is displayed .

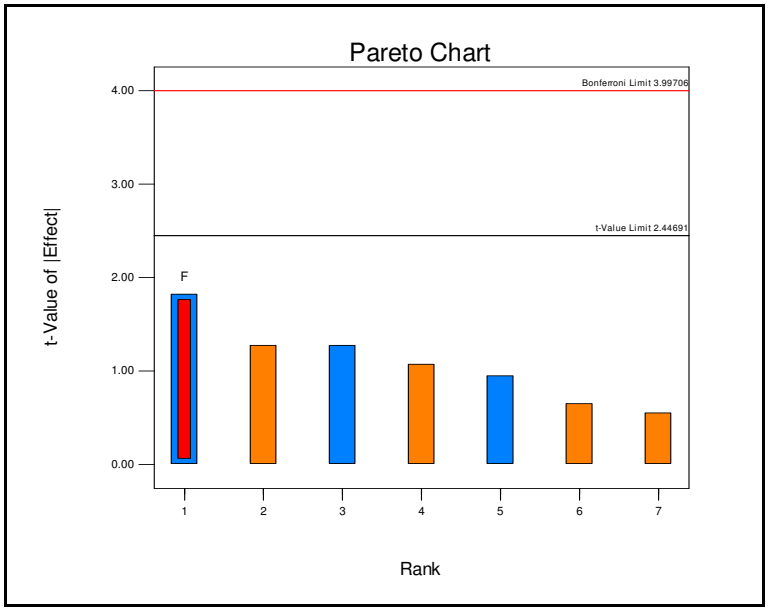

Figure 4.22: Pareto Chart obtained from the 2^{6-3} FFD applied to the Water sample. No factor is significant.

The results obtained from the 2^{6-3} FFD show that the matrix effects do not depend on any particular furnace parameter, except for the Methanol sample, where the T_{Ash} factor was found to be significant.

This leads to two possible conclusions:

1. The sensitivity depends only (or very strictly) on the atomisation process (Section 4.4).
2. As the matrix effects are not generated by the GFAAS experimental procedure, they may be related to a physical part of the instrument, and the graphite tube is the component that interacts directly with the sample (Section 4.5).

4.4 Studying the atomisation process – 2^2 CFD

The atomisation parameters (in particular the atomisation temperature) are the most crucial furnace factors (Section 2.1.2). The sensitivity of a GFAAS analysis depends on the amount of atomised elements and their speed of formation. In fact, the atomisation hold time (Ht_{Atom}) and final temperature (T_{Atom}) were not included in the experimental designs, and the atomisation ramp time (Rt_{Atom}) was excluded after the first screening design (Section 4.2). The experiments described in Section 4.3 proved that the sensitivity might depend *only* on the atomisation process, without a major influence from the other parameters.

To gain insight into the importance of the atomisation process, a 2^2 CFD was proposed. This design only included the atomisation ramp time and final temperature, as the atomisation hold time must be as short as possible. Following the peak shape in Figure 4.12b, it was decided to set the Rt_{Atom} high level at 3 seconds (as opposed to 4 seconds).

This design was used for Methanol, Octanal and Toluene. The design matrix and the results are shown in Table 4.8. The correspondent Pareto charts are shown in Figures 4.23, 4.24 and 4.25 (the ANOVA tables are listed in Appendix A: Table A.10, A.11 and A.12). A centre point was added to the design; the

temperature value of the centre point is the actual atomisation temperature implemented for the industrial application.

Table 4.8: 2^2 CFD for the atomisation parameters investigation (a centre point was added to the design)

Run	Levels	Rt_{Atom} (s)	T_{Atom} (ºC)	Abs		
				Methanol	Octanal	Toluene
1	--	1	2200	0.075	0.106	0.055
2	+-	3	2200	0.063	0.093	0.049
3	-+	1	2600	0.217	0.203	0.148
4	++	3	2600	0.197	0.187	0.135
5	00	2	2400	0.173	0.155	0.093

T_{Atom} is the significant factor. By comparing the results given in Section 4.2 (where Rt_{Atom} was the most significant factor) with Figures 4.23, 4.24 and 4.25, the extent of the importance of T_{Atom} (and the insignificance of the other factors) becomes clear. This is due to the fact that the T_{Atom} low level is below the Rh atomisation temperature. If the low level was higher (e.g. 2400ºC; see Table 4.8) the atomisation temperature would be adequate for each treatment and its significance would be smaller.

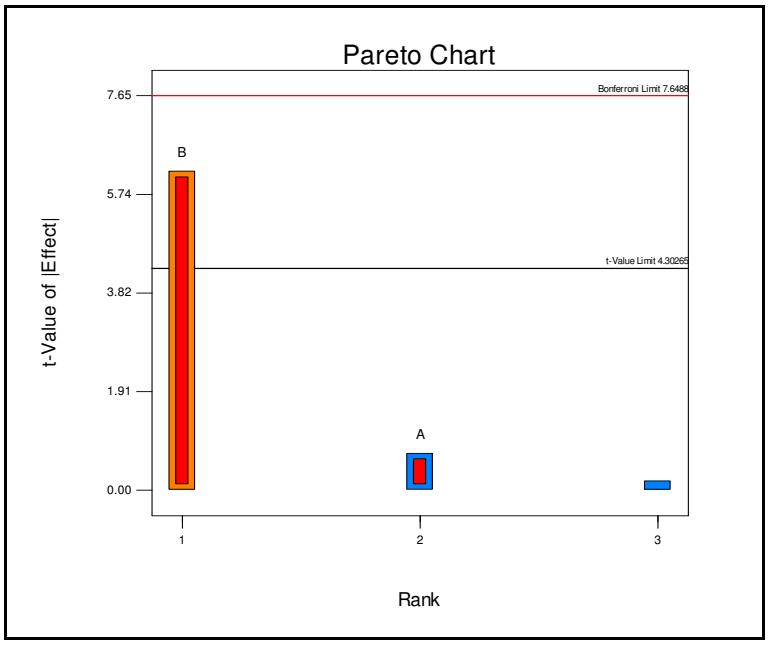

Figure 4.23: Pareto Chart obtained from the 2^2 CFD applied to the Methanol sample. T_{Atom} (B) is a significant factor while Rt_{Atom} (A) is a non-significant factor.

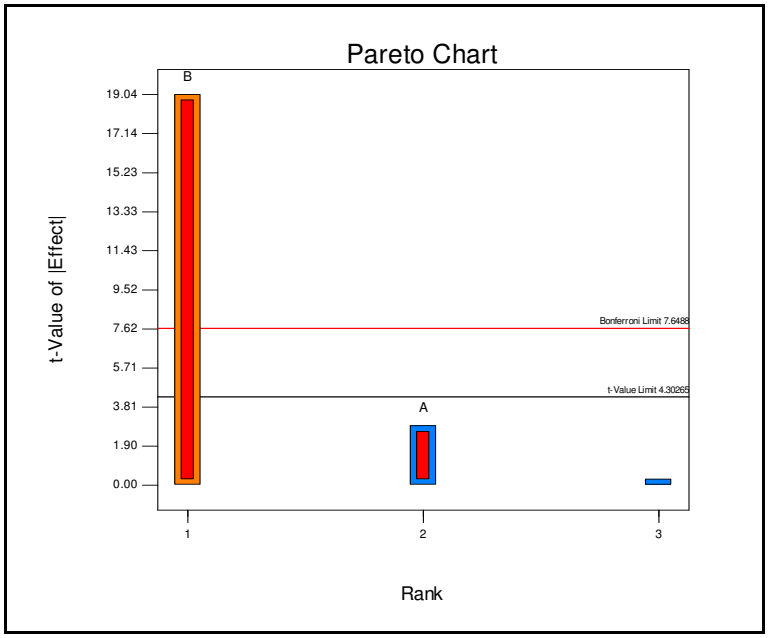

Figure 4.24: Pareto Chart obtained from the 2^2 CFD applied to the Octanal sample. T_{Atom} (B) is a strongly significant factor while Rt_{Atom} (A) is a non-significant factor.

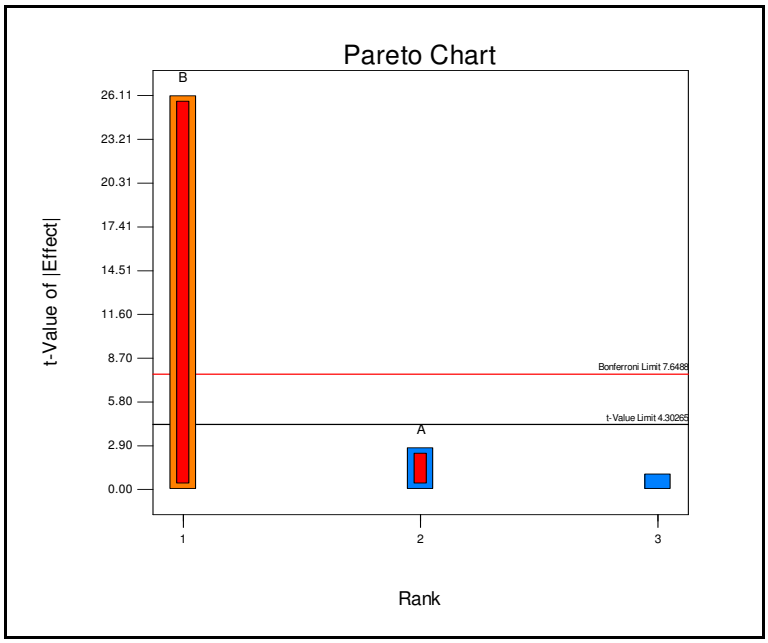

Figure 4.25: Pareto Chart obtained from the 2^2 CFD applied to the Toluene sample. T_{Atom} (B) is a strongly significant factor while Rt_{Atom} (A) is a non-significant factor.

This was also observed by Lobo *et al*. in a study on the optimisation of Cu, Pb, Ni and Cd determination by GFAAS, where the atomisation temperature was varied between 1400 and 1500ºC and it proved to be less important than the sample preparation.[77] On the other hand, while investigating the determination of As in petroleum streams, Cassella *et al*. varied the atomisation temperature between 1900 and 2500ºC observing that this factor has a significant effect on the extent of the absorbance signals. They found that 2600ºC was the optimal value, which is significantly higher than 1900ºC.[26]

By changing the T_{Atom} factor it could be possible to "adjust" the instrumental response when different solvents are employed. However, by using this approach, only the absolute sensitivity of the instrument would be "adjusted" and not the matrix effects.

The results presented in this section could be used to optimise the atomisation parameters. Although this is not part of the scope of this study, an example of this is given in Figures 4.26 and 4.27 (from the Methanol sample). From these 3D graphs one can see that the maximum Abs response is obtained when the atomisation temperature is at the high level and the atomisation ramp time is at the low level (confirming the results presented in Section 4.2).

In the case there were any constrictions/limitations on the parameters, one could use the *Desirability* function to find the best settings within these limits.[94] For example, if the best absorbance must be obtained at the lowest T_{Atom} and Rt_{Atom}, the desirability function can suggest the lowest T_{Atom} and Rt_{Atom} that guarantee good sensitivity. In other words, T_{Atom} and Rt_{Atom} must be minimised while the absorbance must be maximised. Figure 4.28 shows that the highest desirability (0.654; the maximum is 1) is achieved when T_{Atom} is at 2430ºC and Rt_{Atom} is at 0.90 seconds, which should yield an absorbance signal of 0.164 (Figure 4.27 and 4.29) – while the maximum Abs (0.217) is obtained at the highest temperature (Table 4.8). Figure 4.30 shows the 3D surface plot of the desirability, which relates to the 3D surface plot of the Abs response (Figure 4.26 and 4.27).

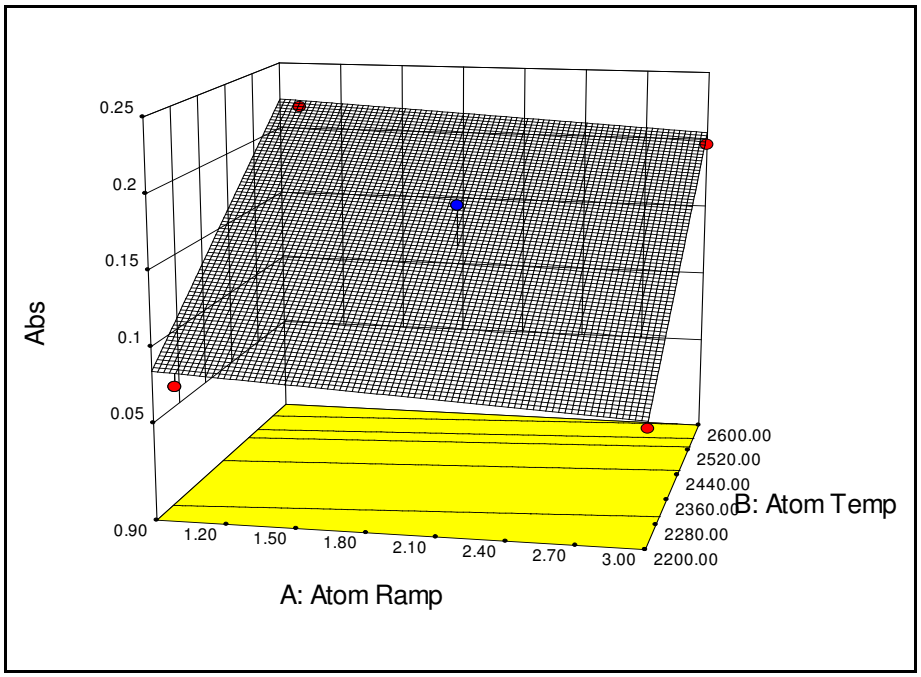

Figure 4.26: 3D surface response obtained from the 2^2 CFD applied to the Methanol sample. The maximum Abs response is obtained when T_{Atom} (B) is high and Rt_{Atom} (A) is low.

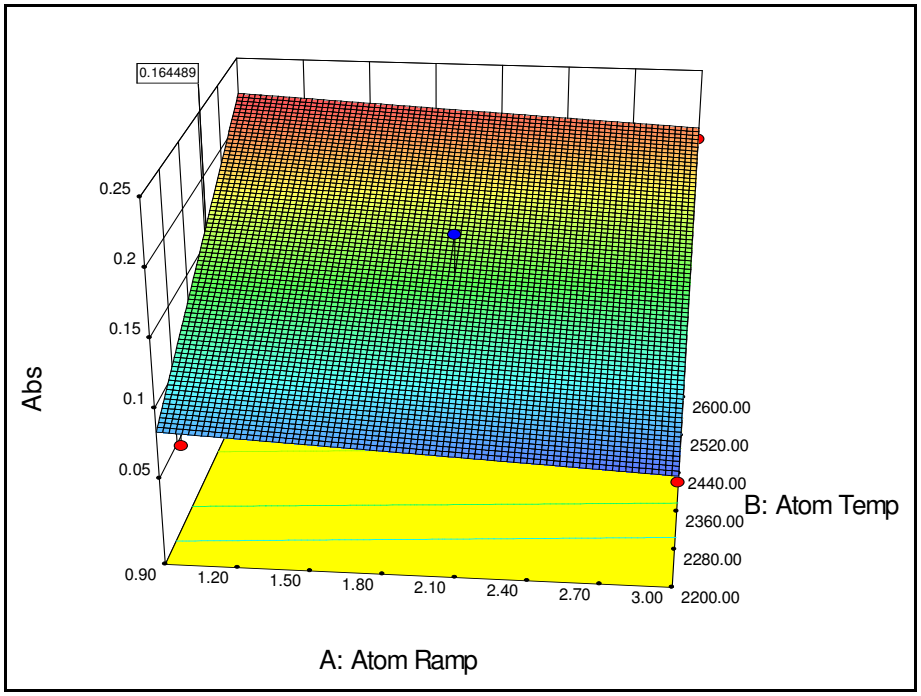

Figure 4.27: Colour 3D surface response obtained from the 2^2 CFD applied to the Methanol sample. The maximum Abs response is obtained when T_{Atom} (B) is high and Rt_{Atom} (A) is low.

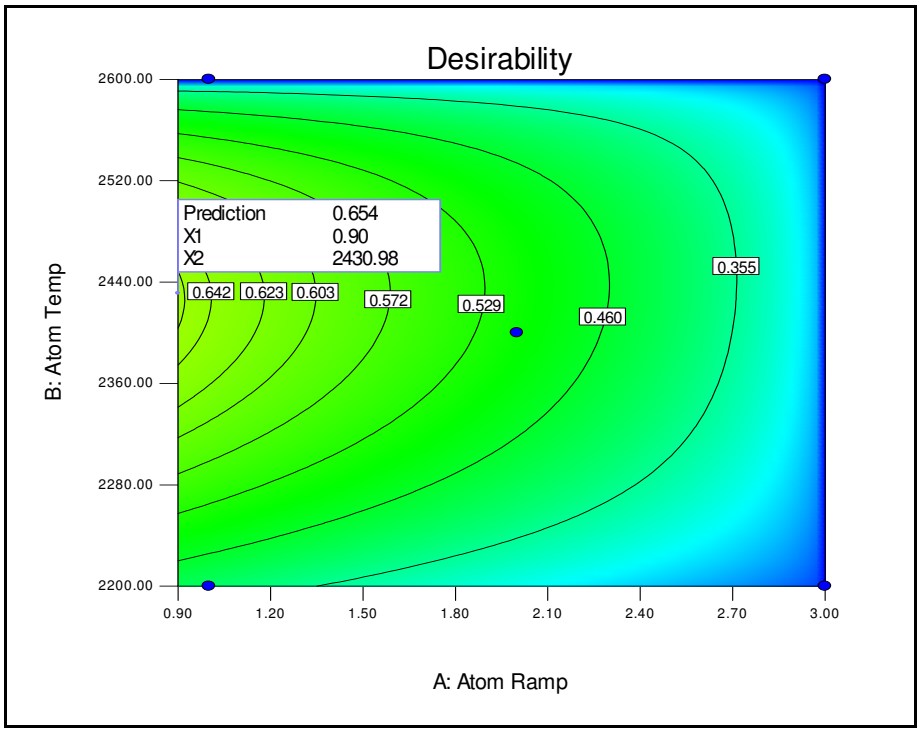

Figure 4.28: Desirability contour plot. The highest desirability (0.654) is predicted when T_{Atom} (B) is at 2430ºC and Rt_{Atom} (A) is at 0.90 seconds.

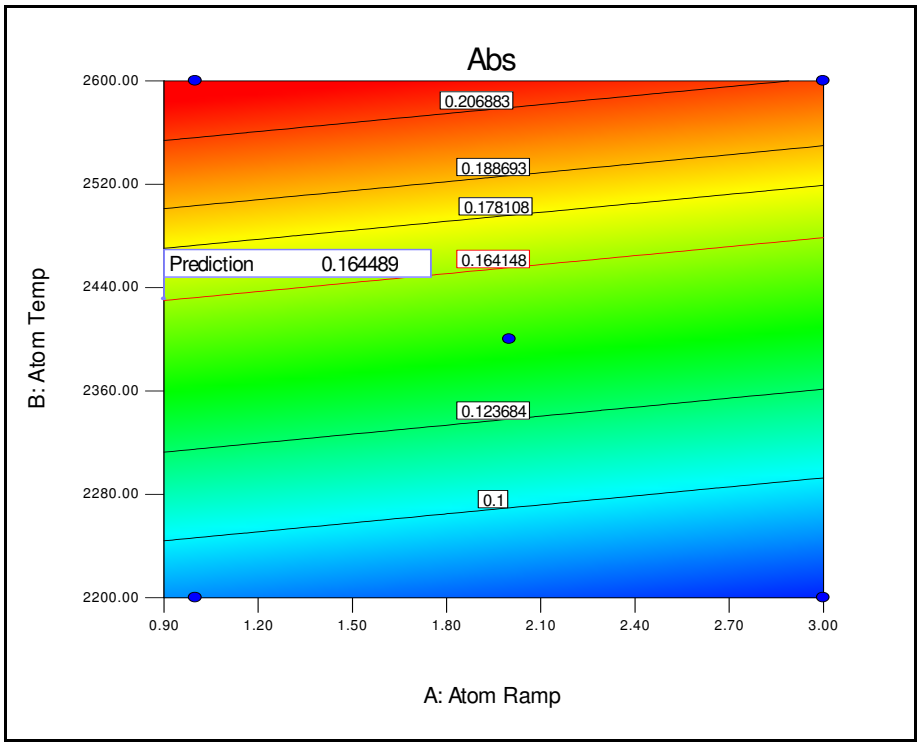

Figure 4.29: Contour plot showing the predicted Abs at the "desired" (suggested) settings for T_{Atom} (B) and Rt_{Atom} (A).

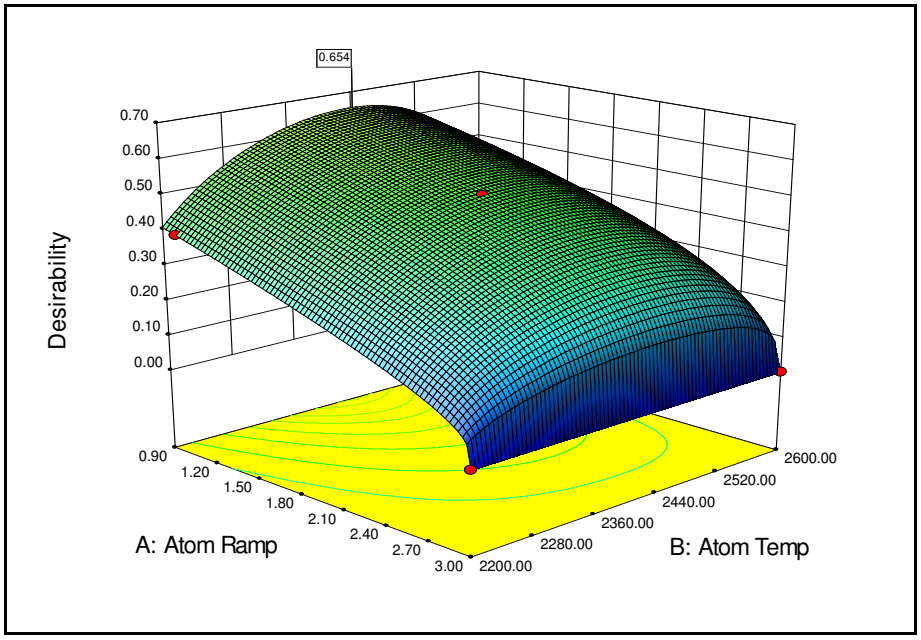

Figure 4.30: Desirability 3D surface response plot. The highest desirability (0.654) is predicted when T_{Atom} (B) is at 2430ºC and Rt_{Atom} (A) is at 0.90 seconds.

Section 4.5 discusses the experiments performed to determine if there are any interactions between the surface of the graphite tube and the sample matrix.

4.5 SEM Analysis

The carbon of graphite tubes plays an important role in GFAAS. This feature has been observed in the past and documented in literature[95]. For example, Grotti *et al.* studied the interferences of Na, K, Ca and Mg on the atomisation of Te and found that when a Ta-tube platform was used, or when the graphite platform was coated with Ni or Pd, no/less interferences were observed.[59] Also, L'vov reported the effect of Ta platforms in terms of the changes produced on the shape of the Abs signal.[96] Similarly, de Amorim *et al.* demonstrated that using a Zr modifier improves the determination of Al by GFAAS and found that, by means of a factorial design, the choice of the modifier is an important factor (simplistically, modifiers can alter the tube surface).[73]

It is possible that the surface of the graphite tubes interacts differently with different solvents, thus generating matrix effects. As a consequence, if different degradation paths are produced, the aspect of graphite surfaces that have been treated with different solvents should be dissimilar.

Three used graphite tubes were subjected to scanning electron microscope (SEM) analysis. The three tubes were treated with Water, Octanal and Hexane. The SEM images are shown in Figures 4.31, 4.32 and 4.33 (600x magnification), for Water, Octanal and Hexane, respectively.

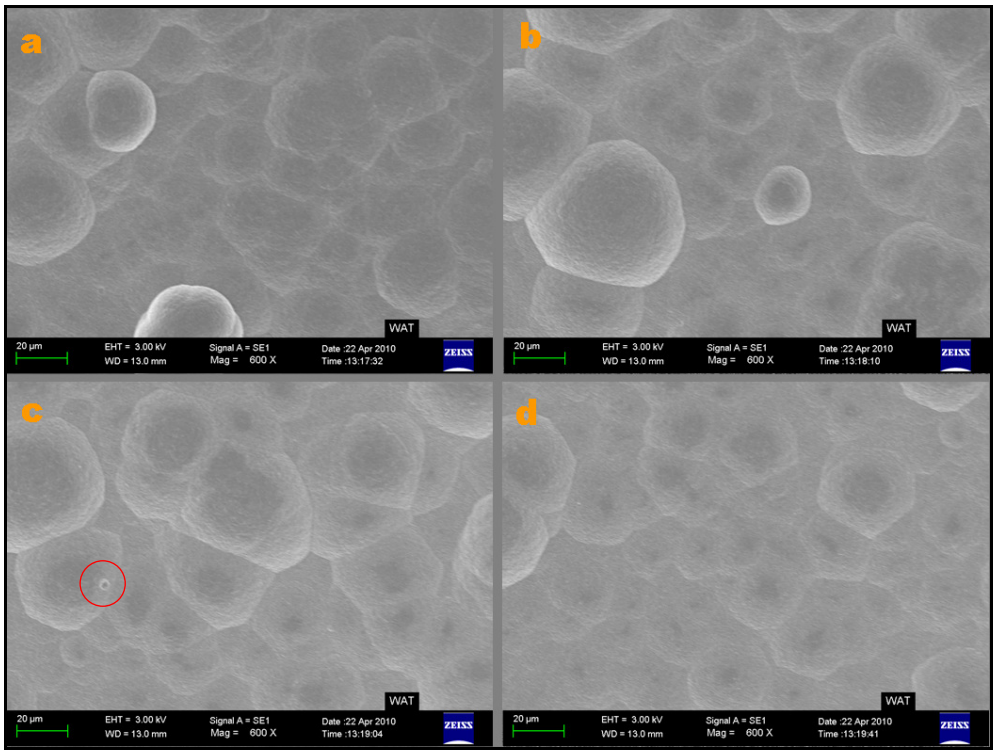

Figure 4.31: SEM image of graphite tube treated with Water (600x magnification). The different pictures (a, b, c and d) were taken from different sections of the tube. The red circle indicates a blister.

There is no remarkable difference between the SEM images of the three tubes. Nonetheless, entities that can be defined as *nodules*[97] display different sizes and distributions. The sizes of the nodules are identified as small, medium and large. Furthermore, small marks can be distinguished that could be referred as *blisters*.[98]

The tube treated with Water does not display a uniform surface (Figure 4.31). The distribution of the nodules is irregular and only one distinct blister is observed (Figure 4.31c).

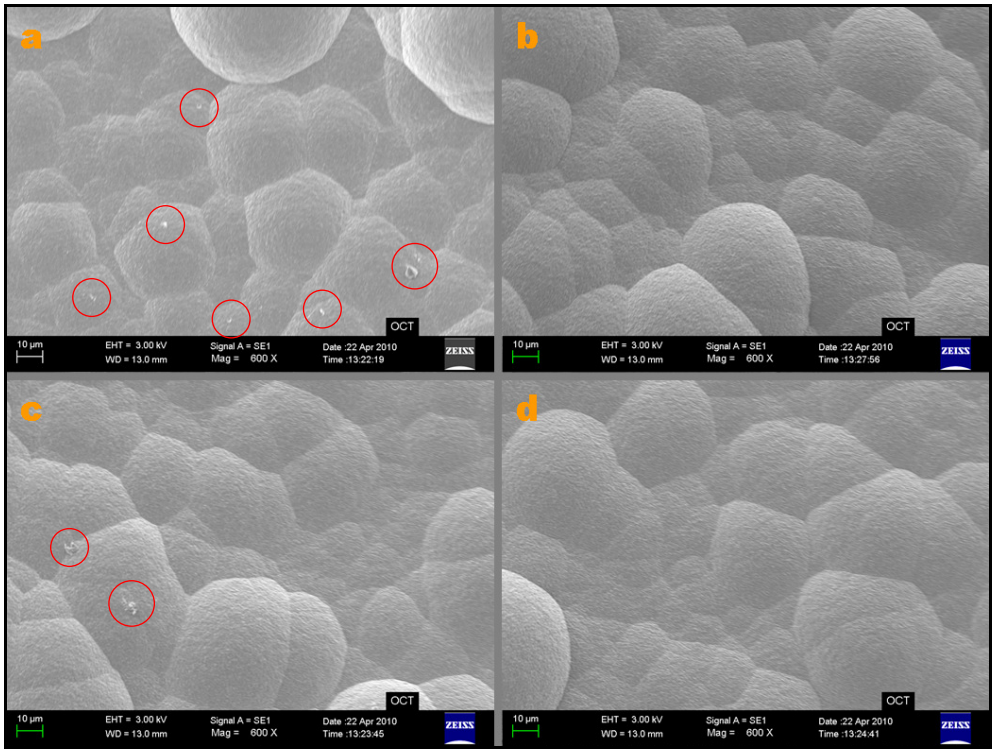

Figure 4.32: SEM image of graphite tube treated with Octanal (600x magnification). The different pictures (a, b, c and d) were taken from different sections of the tube. The red circles indicate the blisters.

Figure 4.32 (Octanal treated tube) shows a more uniform distribution of medium sized nodules. This picture is characterised by a significant amount of blisters, especially in the region depicted in Figure 4.32a. This could indicate that different regions of a tube are subjected to different conditions.

Figure 4.33 represents the tube treated with Hexane. This surface looks more flat than the others due to the small sized, closely packed nodules. Few large nodules are visible in Figures 4.33b and 4.33d. Small blisters can be distinguished in this Figure, except for a very large one in the region captured in Figure 4.33d. The differences in this region could indicate that this part of tube

was subjected to harsher conditions (such as higher temperatures or faster temperature ramps) than the other areas.

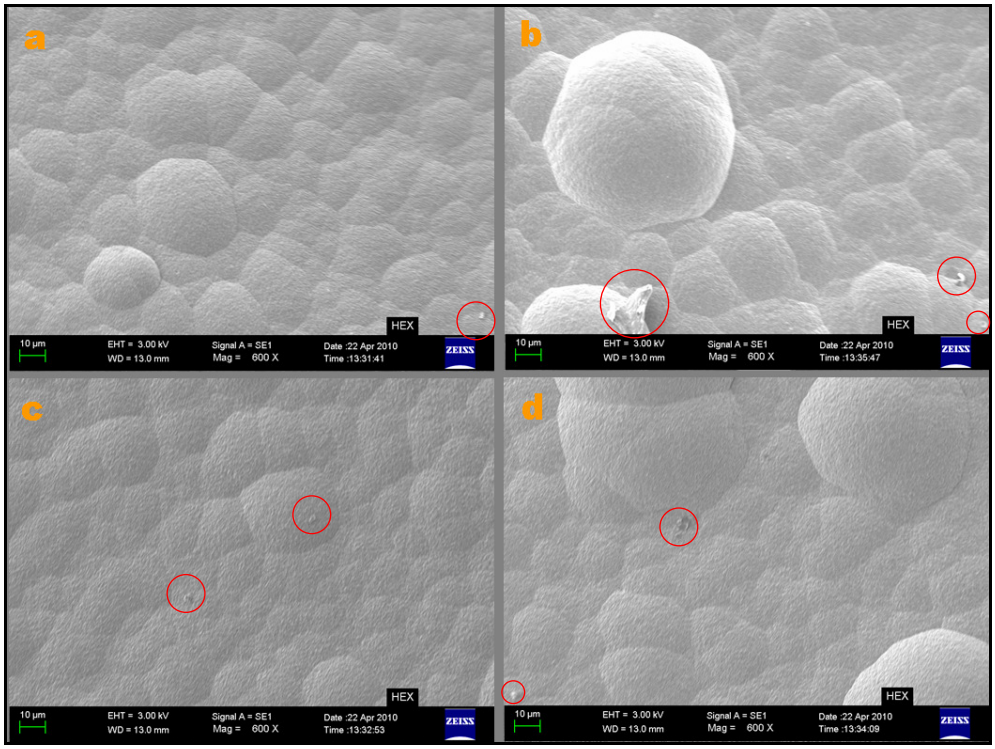

Figure 4.33: SEM image of graphite tube treated with Hexane (600x magnification). The different pictures (a, b, c and d) were taken from different sections of the tube. The red circles indicate the blisters.

The presence of blisters has already been observed by Welz *et al.* during their studies on the lanthanum effects on graphite surfaces.[98] In a related investigation, Ortner *et al.* observed that the nodules form only at the hottest sections of the tube, where gaseous carbon condenses when the temperature decreases.[97]

The subtle differences between Figures 4.31, 4.32 and 4.33 cannot prove (at least at this stage) that the matrix effects are caused by the interactions between the graphite tube surface and the different solvents. However, if these results are considered together with the knowledge that the graphite carbon plays an important role during the furnace cycle, it might be useful to pursue this kind of studies. Different GFAAS experiments followed by SEM analyses could

proof/disproof that the matrix effects depend on the interactions mentioned above. Examples of such experiments are given in Chapter 5.

Section 4.6 describes how the results obtained from the experimental design in Section 4.3 are analysed by PCA.

4.6 PCA input

The experimental results discussed in this Chapter can be considered as a dataset that contains some "hidden" information. To further interpret this information a PCA was performed on the results obtained from the 2^{6-3} FFD discussed in Section 4.3. These results should contain the most relevant knowledge. As discussed above, the results obtained from the 2^{7-4} FFD are not reliable (Section 4.2) while the results obtained from the 2^2 CFD that considered the atomisation step carry an "obvious" knowledge (Section 4.4).

In PCA the information of the dataset is explained by the PCs (Section 2.2.1). The significance of each PC (i.e. how much information they hold) is expressed in terms of *eigenvalues* and it is represented graphically by the total explained variance plot. The PCs with a very small eigenvalue (generally the higher PCs) carry the instrumental noise only.[63,65]

A PCA was performed on the dataset in Table 4.9, which lists the results obtained from each treatment (T1-T8). Figure 4.34 shows the total explained variance plot obtained from this analysis. Figure 4.35 shows the score plot and Figure 4.36 shows the loading plot.

Table 4.9: Dataset analysed by PCA.

Objects	Variables							
	T1	**T2**	**T3**	**T4**	**T5**	**T6**	**T7**	**T8**
Water	0.182	0.103	0.197	0.212	0.191	0.189	0.184	0.174
Methanol	0.197	0.198	0.213	0.203	0.21	0.213	0.198	0.205
Octanal	0.210	0.187	0.220	0.224	0.227	0.269	0.208	0.204
Toluene	0.143	0.149	0.164	0.162	0.162	0.162	0.161	0.128
Hexane	0.049	0.009	0.022	0.027	0.025	0.028	0.011	0.025

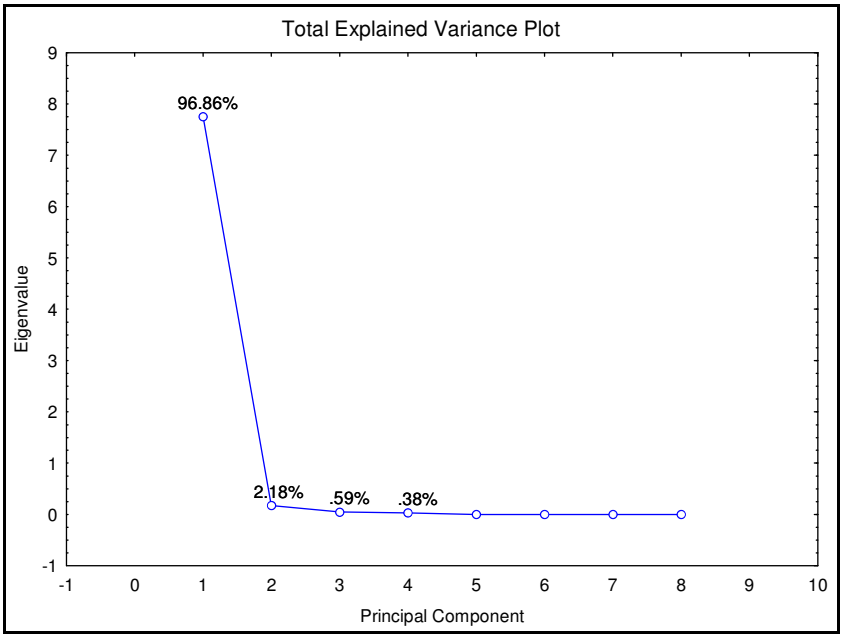

Figure 4.34: Total explained variance plot from the dataset in Table 4.9. PC1 explains 96.89% of the information hidden in the data set.

Figure 4.34 shows that PC1 explains 96.86% of the information while PC2 explains only 2.18%. PC3 and PC4 are related to instrumental noise.

The score plot in Figure 4.35 reveals that PC1 corresponds to the sensitivity of the GFAAS analysis. If the projections of the objects (solvents) on the *x*-axis are considered, the order, from left to right, would be Octanal, Methanol, Water, Toluene and Hexane, which reflects the Absorbance trend shown by the averages in Table 4.7. PC2 might symbolise the solvent nature, with Water on the bottom side of the plot and the organic solvents on the top. However, 2.18% is a small eigenvalue and also PC2 might be associated with noise.

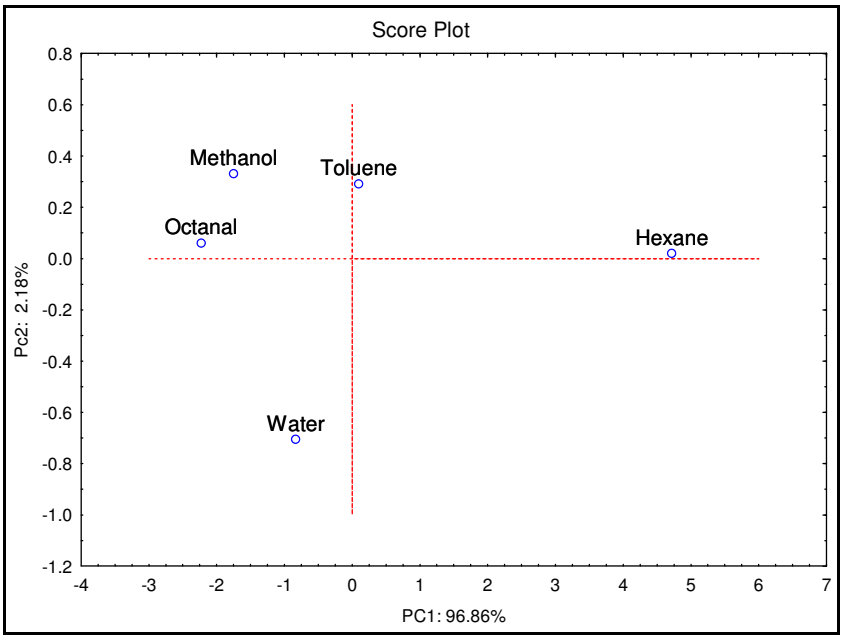

Figure 4.35: Score plot from the dataset in Table 4.9. PC1 is on the x-axis and PC2 on the y-axis.

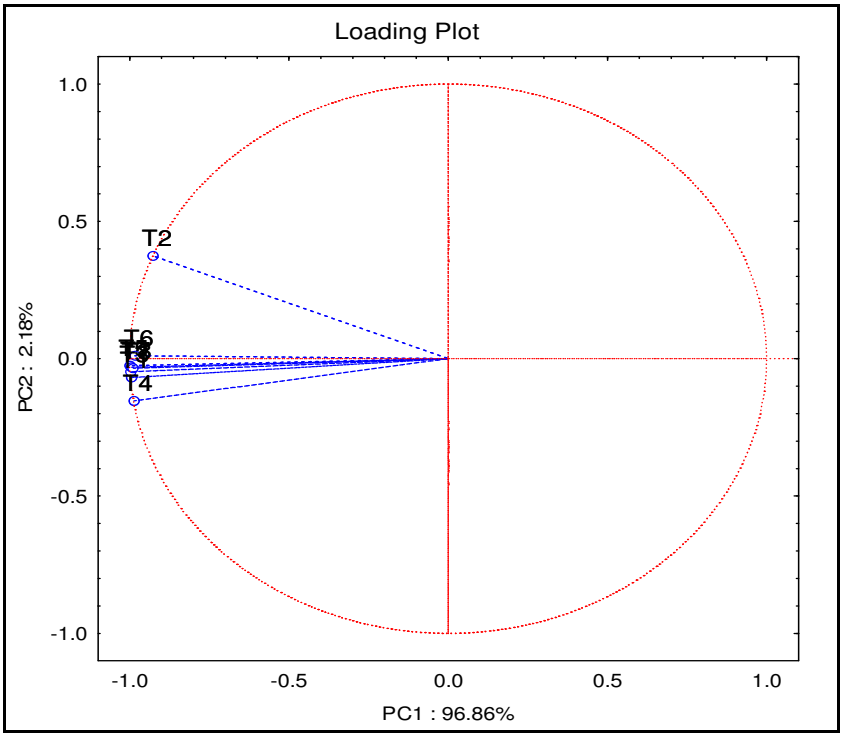

Figure 4.36: Loading plot from the dataset in Table 4.9. PC1 is on the x-axis and PC2 on the y-axis.

The loading plot in Figure 4.36 confirms that all treatments (and the sensitivity) are strongly associated with PC1. T2 could be considered an outlier due to the

separation from the main group. This conclusion could be verified by the results obtained from the second treatment (Table 4.9) which seem low for Octanal, Water and Hexane. In contrast, the distance between T2 and the other treatments is not truly significant and the angle is smaller than 90º – a right angle would indicate that T2 is not related to the other treatments, supporting the hypothesis of an outlier. Furthermore, if we compare the average of the treatment results with the average calculated without including T2, only one significant discrepancy is observed (i.e. Water; Table 4.10).

Table 4.10: Averages of the treatment results (Abs).

Sample	Treatment average	Treatment average without T2
Water	0.179	0.190
Methanol	0.205	0.206
Octanal	0.219	0.223
Toluene	0.154	0.155
Hexane	0.025	0.027

Table 4.11: PCA Dataset with the inclusion of physical constants: molar mass (MM, g/mol), density (D, Kg/m^3), melting point (MP, ºC), boiling point (BP, ºC), surface tension (ST, mN/m) and viscosity (V, mPa).[99-101]

Objects	Variables						
	MM	D	MP	BP	ST	V	T1
Water	18.02	1000	0	100	71.99	0.89	0.182
Methanol	32.04	791.3	-97	64.7	22.07	0.544	0.197
Octanal	128.21	820	13	171	26.8	1.2	0.21
Toluene	92.14	866.9	-93	110.6	27.73	0.56	0.143
Hexane	86.18	654.8	-95	69	17.89	0.252	0.049

	T2	T3	T4	T5	T6	T7	T8
Water	0.103	0.197	0.212	0.191	0.189	0.184	0.174
Methanol	0.198	0.213	0.203	0.210	0.213	0.198	0.205
Octanal	0.187	0.220	0.224	0.227	0.269	0.208	0.204
Toluene	0.149	0.164	0.162	0.162	0.162	0.161	0.128
Hexane	0.009	0.022	0.027	0.025	0.028	0.011	0.025

The dataset in Table 4.9 does not contain enough data to classify the solvents. It does not include any information on the nature of the solvents and the only explained knowledge is the sensitivity trend of the experiments.

Table 4.11 include some physical constants: molar mass (MM, g/mol), density (D, Kg/m^3), melting point (MP, ºC), boiling point (BP, ºC), surface tension (ST, mN/m) and viscosity (V, mPa).[99-101] The computed total explained variance plot is shown in Figure 4.37. Some score and loading plots are shown from Figure 4.38 to Figure 4.41. This new dataset includes variables characterised by different units. Hence, the data needs to be standardised prior the analysis, using the *correlation matrix* instead of the *covariance matrix*.

The loading plot in Figure 4.38 shows that PC1 is still associated with the treatment results (i.e. the sensitivity) but it is also related to melting point and viscosity. As in Figure 4.35, a sensitivity trend along PC1 is observable in Figure 39. However, Water is positioned before Methanol, although Methanol is more sensitive. This is due to the contributions of melting point and viscosity to the data reduction; hence, the phenomenon represented by PC1 is not just the sensitivity of the analysis, but a feature related to sensitivity, melting point and viscosity (which could not be determined with the dataset in Table 4.9). PC1 also exhibits a viscosity trend, from Octanal (highest viscosity) to Hexane (lowest viscosity); with the exception that Methanol is placed before Toluene (Methanol is less viscous than Toluene). Water and Octanal, on the left hand side of the plot, are also the objects with the highest melting points. This leads to the possibility that the sensitivity (and possibly the matrix effects) may be linked to the viscosity and melting point of the samples.

Figure 4.38 also shows that the molar mass is strongly correlated to PC2. In fact, in Figure 4.39, PC2 orders the solvents from the heaviest (Octanal) to the lightest one (Water). This means that the molar mass is not correlated to PC1 (MM is perpendicular to PC1) and it probably does not influence the sensitivity of GFAAS analyses. Due to the chemical relationships between molar mass and boiling point, it is possible that BP (at halfway between PC1 and PC2 in Figure 4.38) is also not correlated to PC1.

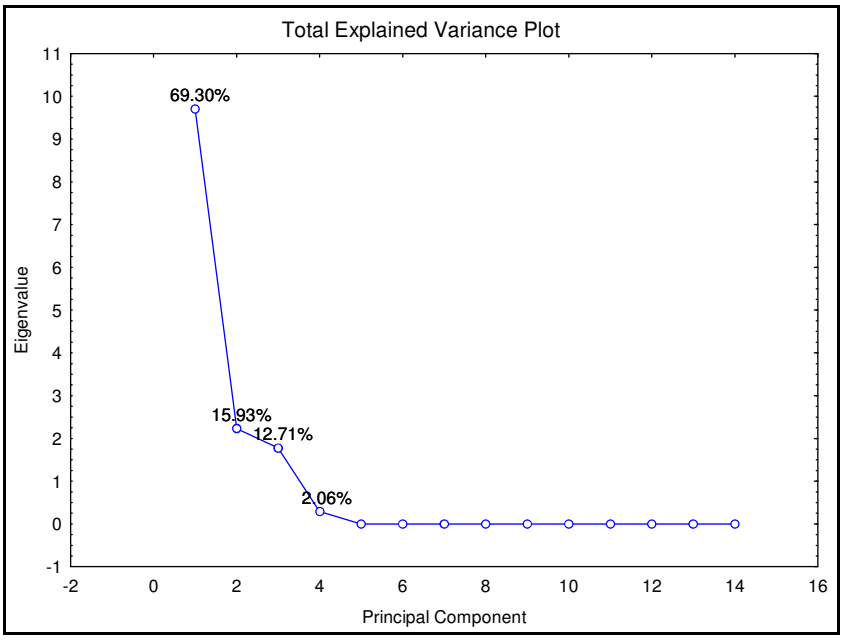

Figure 4.37: Total explained variance plot from the dataset in Table 4.11. PC1 explains 69.30% of the information hidden in the data set, while PC2, PC3 and PC4 holds 15.93%, 12.71% and 2.06%, respectively..

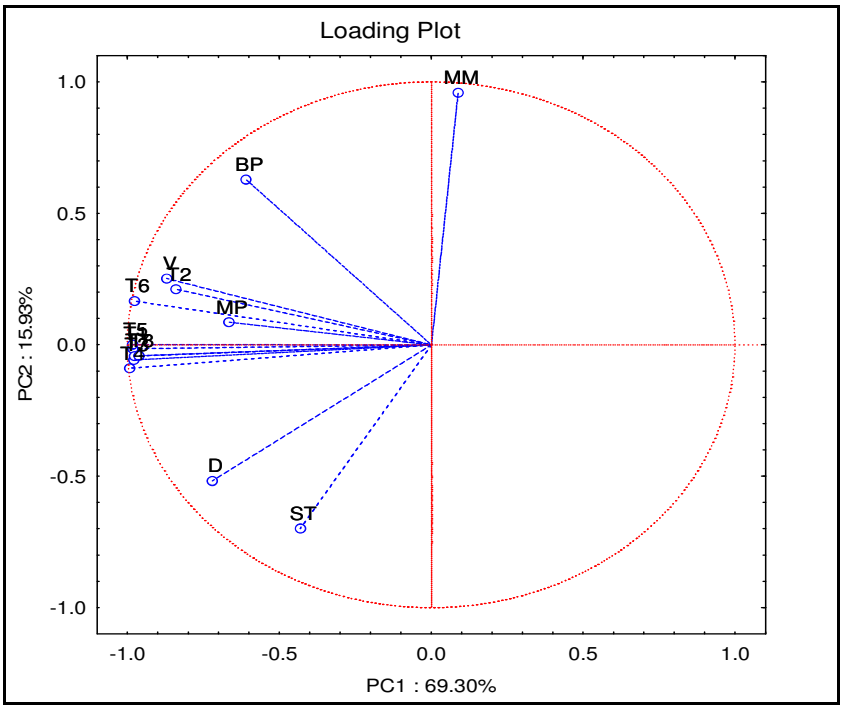

Figure 4.38: Loading plot from the dataset in Table 4.11. PC1 is on the x-axis and PC2 on the y-axis.

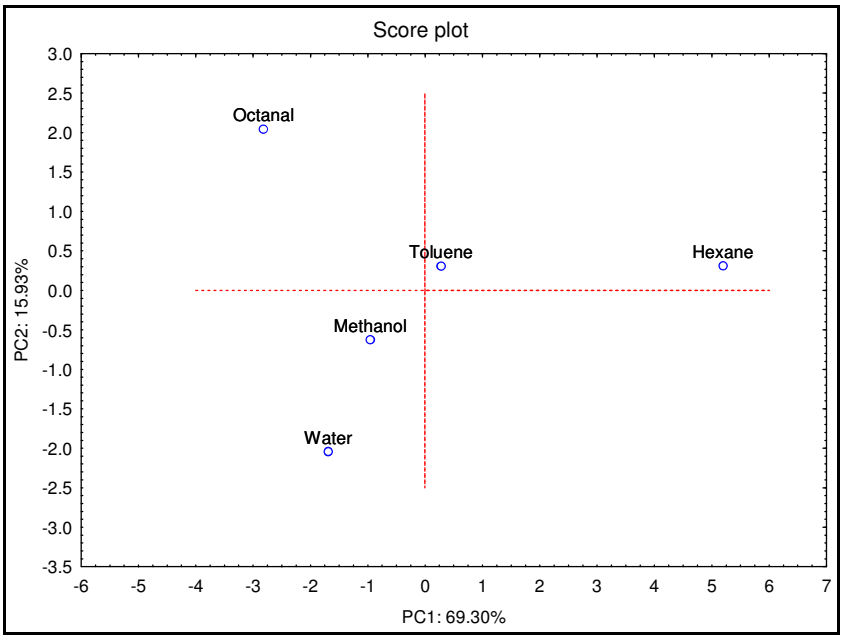

Figure 4.39: Score plot from the dataset in Table 4.11. PC1 is on the x-axis and PC2 on the y-axis.

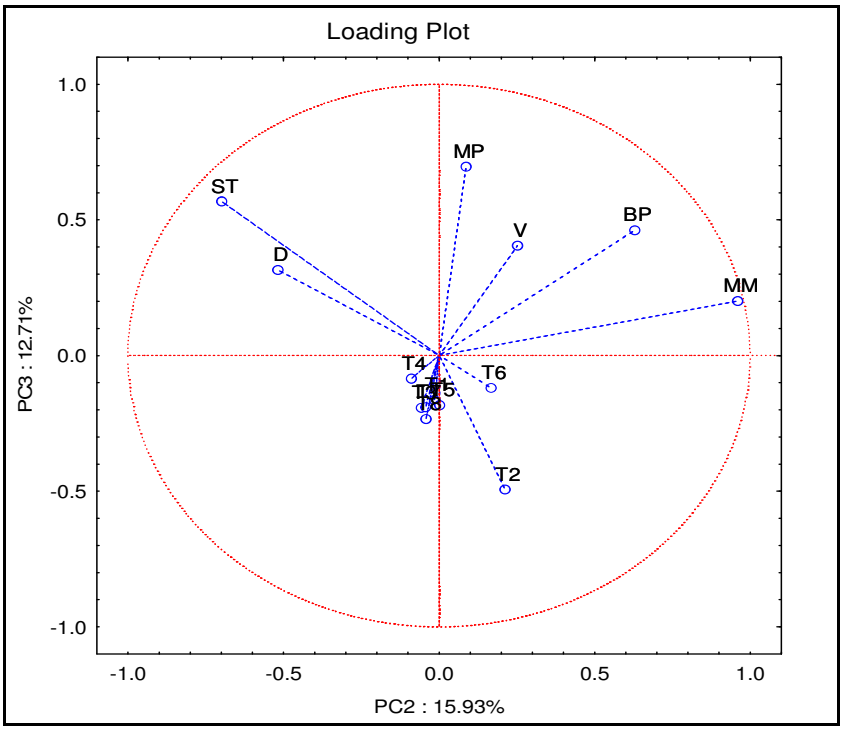

Figure 4.40: Loading plot from the dataset in Table 4.11. PC2 is on the x-axis and PC3 on the y-axis.

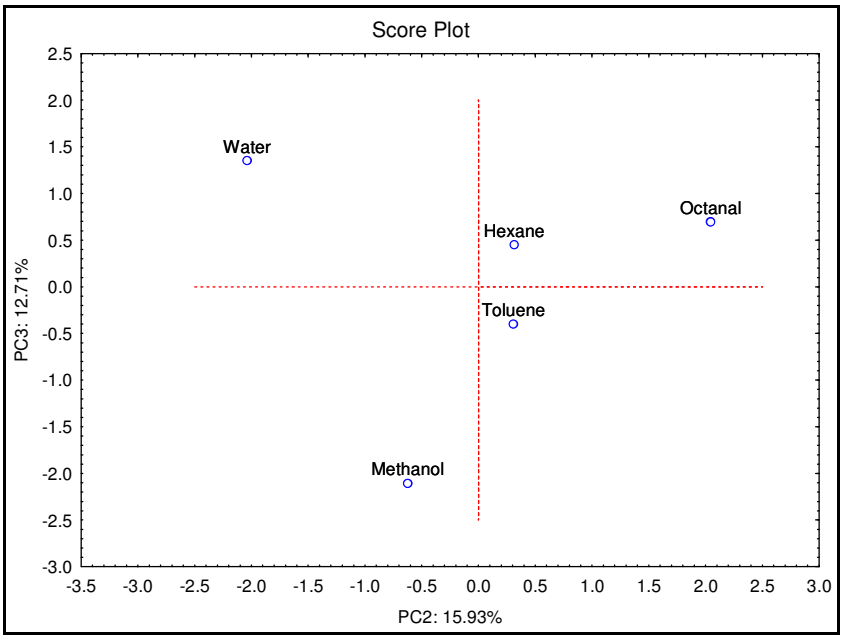

Figure 4.41: Score plot from the dataset in Table 4.11. PC2 is on the x-axis and PC3 on the y-axis.

PC3 contains 12.71% of the information of the dataset and cannot be associated with noise. Nonetheless, Figure 4.40 shows that there is no variable directly linked to PC3 (except a small correlation with the melting point), and no clear phenomenon can be identified by it. The inclusion of other variables in the dataset might reveal the meaning of PC3.

The majority of the knowledge (69.30%, explained by PC1) contained in Table 4.11 consists of the sensitivity exhibited by the different solvents. Figures 4.38 and 4.39 show that viscosity and melting point are correlated to PC1 and may influence the sensitivity of the analysis – although Castillo *et al.* reported that no correlation was found between sensitivity and physical properties.[102] 15.93% of this knowledge says that the molar mass is not correlated with the sensitivity of the GFAAS analysis (it is perpendicular to PC1); this information could be useful to understand the nature of matrix effects. 12.71% (PC3) of the information cannot be identified as the variables employed for the data reduction are not directly linked to PC3.

5 CONCLUSION AND RECOMMENDATIONS FOR FUTURE WORK

Matrix effects were observed in Rh analyses by GFAAS. Due to these effects the analysis was based on a standard addition calibration method, which is not efficient from a commercial point of view (Section 1.1). Experimental design (Section 2.3) was proposed to investigate several GFAAS parameters and assess their role in the generation of these matrix effects.

A preliminary set of experiments was performed to examine the ashing step, where matrix-graphite (carbon) interactions are supposed to occur. From these experiments it was observed that the ashing step alone is not significant in relation to sensitivity and matrix effects (Section 4.1).

A 2^{7-4} FFD was proposed to investigate seven furnace factors: Rt_{Tr}, Ht_{Tr}, T_{Tr}, Rt_{Ash}, Ht_{Ash}, T_{Ash} and Rt_{Atom} (Section 4.2). The other furnace parameters (Rt_{Dry}, Ht_{Dry}, T_{Dry}, Ht_{Atomr} and T_{Atom}) were considered not suitable for a factorial design due to their role in GFAAS. The results of this DoE showed that the only significant factor was Rt_{Atom}. This is due to the important role that this factor has on the generation of free atoms (hence on the sensitivity). It was decided to exclude this factor from the design and investigate the six remaining factors with a 2^{6-3} FFD (Section 4.3).

The results obtained from the 2^{6-3} FFD showed that no factor can be considered responsible for the matrix effects. Except for the Methanol sample (where T_{Ash} was significant), the sensitivity of other samples did not depend on any specific factor (no significant factor was found). Hence, the sensitivity depends only (or very strictly) on the atomisation process. Alternatively, the matrix effects may be related to interactions between sample matrix and graphite tube surface.

A 2^2 CFD was performed to assess the importance of the atomisation process (Section 4.4). The factors investigated were T_{Atom} and Rt_{Atom}. This design was applied to the Methanol, Octanal and Toluene samples. The results showed that T_{Atom} is the most significant factor, as is generally expected in GFAAS. This result can only be used to optimise the atomisation step (ignoring the matrix effects) and an example is given in Section 4.4.

It is possible that the surface of the graphite tubes experiences different interactions with different solvents, thus generating matrix effects. Hence, three used graphite tubes (treated with Water, Octanal and Hexane) were analysed by SEM to asses whether the tubes were subjected to different degradation paths (Section 4.5). The differences between these SEM images are subtle and cannot prove that the matrix effects depend on the interactions between the graphite surface and the matrix. However, it is known that the graphite carbon plays an important role during the furnace cycle; thus, it might be useful to perform more GFAAS experiments followed by SEM analyses. Examples of such experiments are listed below:

- Tube treated with more solvents or mixtures of solvents;
- Addition of modifiers prior to solvent treatment (to alter the surface of the graphite tube);
- Tube subjected to an incomplete furnace cycle that does not advance to the atomisation step, to assess if the high atomisation temperature removes the evidences of any interaction between tube surface and solvents. In this way only the ashing step would be investigated.

Unfortunately, this kind of research is far from time or cost effectiveness. GFAAS is generally a slow technique and a considerable amount of time is required if conclusive results need to be achieved. Furthermore, one graphite tube must be employed and destroyed (for SEM analysis) for each experiment, limiting the number of experiments because of cost implications.

To gain more knowledge from the results, a PCA was performed on a dataset composed of the GFAAS results and some physical properties of the solvents analysed (Section 4.6). PC1 (which carries 69.30% of the knowledge hidden in

the dataset) is correlated to sensitivity, viscosity and melting point. PC2 (15.93%) reveals that the molar mass is not linked to PC1. No variable was directly/strongly correlated to PC3 (12.71%), probably because more variables should have been used for the data reduction.

A summary of these results is given below:

- The experimental designs in Sections 4.1, 4.2, 4.3 and 4.4 revealed that the sensitivity depends mostly on the atomisation temperature and that the matrix effects are not linked to any furnace parameter.
- The SEM analyses described in Section 4.5 showed that the graphite tube surface might react in different ways with different solvents/matrices. However, this does not prove/disprove that the matrix effects are based on these interactions and more GFAAS-SEM experiments could fill this knowledge gap.
- From the PCA in Section 4.6 it was found that the sensitivity may depend on the viscosity and melting point of the solvents and it is probably independent of the molar mass. However, some of the knowledge hidden in the dataset is still not known and a data reduction based on more variables might provide this information.

In line with these results, the next stage of this research should be the investigation of the link between tube surface-matrix interactions and physical properties. As the sensitivity may depend on viscosity and melting point, a variety of solvents that cover a wide range of these physical properties should be analysed by GFAAS. In addition, the surface of the used tubes should be examined by SEM and the GFAAS results should be organised in a dataset for a PCA. This dataset should include various variables to uncover most of the knowledge and asses if any other factor is correlated to the sensitivity. This approach might uncover the actual causes of the matrix effects. From these causes, one could find the best way to correct or compensate for these interfering effects.

Another approach could be the prevention of matrix effects by means of instrumental adjustments. Najafi *et al.* proposed to separate the analyte from the matrix by electrodeposition, with the electrolytic process occurring inside the graphite tube prior the furnace cycle.[30] Matrix separation/destruction can also be achieved with the aid of automated systems based on flow injection analysis (FIA) that pre-treat the sample before the analysis.[103] Such systems can prevent the occurrence of matrix effects, but they need more complex instrumentations. Higher complexity usually involves more parameters that need to be optimised and this can lead to other limitations or interference sources.

From an industrial point of view, as the GFAAS parameters cannot compensate for the matrix effects, standard additions remain the preferred mode of operation as it accounts for these effects *in-situ*. Although this analytical method involves higher costs and lengthy laboratory analyses, its industrial application is fully justified with samples characterised by complex matrices.

REFERENCES

1. Fermi, E. "Enrico Fermi quotes". *Thinkexist. com* **2010**, viewed on 18 May 2010, http://en.thinkexist.com/quotes/enrico_fermi/

2. Lazzaroni, R.; Settambolo, R.; Alagona, G.; Ghio, C. Investigation of alkyl metal intermediate formation in the rhodium-catalyzed hydroformylation: Experimental and theoretical approaches. *Coordination Chemistry Reviews* **2010,** *254*, 696-706.

3. Shaharun, M. S.; Dutta, B. K.; Mukhtar, H.; Maitra, S. Hydroformylation of 1-octene using rhodium-phosphite catalyst in a thermomorphic solvent system. *Chemical Engineering Science* **2010,** (65), 273-281.

4. Mizner, W. "Wilson Mizner quotes". *Thinkexist. com* **2010**, viewed on 22 July 2010, http://thinkexist.com/quotation/when_you_take_stuff_from_one_writer_it-s/8433.html

5. Dewey, J. "John Dewey Quotes". *BrainyQuote* **2010**, viewed on 22 July 2010, http://www.brainyquote.com/quotes/authors/j/john_dewey.html

6. Dean, J. R. *Atomic Absorption and Plasma Spectroscopy;* Second ed.; John Wiley and Sons: 1997.

7. Lajunen, L. H. J. *Spectrochemical analysis by atomic absorption and emission;* The Royal Society of Chemistry: 1992.

8. Andrade-Garda, J. M. *Basic chemometrics techniques in atomic spectroscopy;* The Royal Society of Chemistry: 2009; Vol. 10.

9. Krasowski, J. A.; Copeland, T. R. Matrix interferences in furnace atomic absorption spectrometry. *Analytical Chemistry* **1979,** *51* (11), 1843-1849.

10. L'vov, B. V. A personal view of the evolution of graphite furnace atomic absorption spectrometry. *Analytical Chemistry* **1991,** *63*, 924 A-931 A.

11. Grégoire, D. C.; Chakrabarti, C. L. Atomization from a platform in graphite furnace atomic absorption spectrometry. *Analytical Chemistry* **1977,** *49* (13), 2018-2023.

12. Pereira-Filho, E. R.; Pérez, C. A.; Poppi, R. J.; Arruda, M. A. Z. Metals distribution and investigation of L'vov platform surface using principal component analysis, micro synchroton radiation X-ray fluorescence spectrometry and scanning electron microscopy after the determination of Al in a milk slurry sample. *Spectrochimica Acta Part B* **2002,** *57*, 1259-1276.

13. Wegscheider, W.; Jancár, L.; Michaelis, M. R. A.; Phe, M. T.; Ortner, H. M. Sensitivity estimation by the anaysis of peak shapes in graphite furnace atomic absorption spectrometry. *Chemometrics and Intelligent laboratory Systems* **1990,** *7*, 281-293.

14. Benzo, Z.; Araujo, P. W.; Sierraalta, A.; Ruette, F. Experimental and theoretical studies of the factors that influence the determination of molybdenum by electrothermal atomic absorption spectroscopy. *Analytical Chemistry* **1993,** *65*, 1107-1113.

15. Bulska, E.; Piascik, M.; Katskov, D.; Darangwa, N.; Grotti, M. Investigation of aging processes of graphite tubes modified with iridium and rhodium used for atomic spectrometry. *Spectrochimica Acta part B* **2007,** *62*, 1195-1202.

16. Jackson, K. W. Electrothermal atomic absorption spectrometry and related techniques. *Analytical Chemistry* **2000,** *72*, 159R-167R.

17. Cruz, R. B.; Van Loon, J. C. A critical study of the application of graphite-furnace non-flame atomic absorption spectrometry to the determination of trace base metals in complex heavy-matrix sample solutions. *Analytica Chimica Acta* **1974,** *72*, 231-243.

18. Rojas, F. S.; Ojeda, C. B.; Pavon J.M.C. Experimental design in the optimization of a microwave acid digestion procedure for the determination of metals in biomorphic ceramic samples by inductively coupled plasma mass spectrometry and atomic absorption spectrometry. *Microchemical Journal* **2010,** *94*, 7-13.

19. Grotti, M.; Gnecco, C.; Bonfiglioli F. Multivariate quantification of spectroscopic interferences caused by sodium, calcium, chlorine and sulfur in inductively coupled plasma mass spectrometry. *Journal of Analytical Atomic Spectrometry* **1999,** *14*, 1171-1175.

20. Chudzinska, M.; Baralkiewicz, D. Estimation of honey authenticity by multielements characteristics using inductively coupled plasma-mass spectrometry (ICP-MS) combined with chemometrics. *Food and Chemical Toxicology* **2010,** *48*, 284-290.

21. Ražic, S.; Đogo, S. Determination of chromium in Mentha piperita L. and soil by graphite furnace atomic absorption spectrometry after sequential extraction and microwave-assisted acid digestion to assess potential bioavailability. *Chemosphere* **2010,** *78*, 451-456.

22. Lepri, F. G.; Welz, B.; Dessuy, M. B.; Vale, M. G. R.; Bohrer, D.; de Loos-Vollebregt, M. T. C.; Dong Huang, M.; Becker-Ross, H. Investigation of the feasibility to use Zeeman-effect background correction for the graphite furnace determination of phosphorous using high-resolution continuum source atomic absorption spectrometry as a diagnostic tool. *Spectrochimica Acta Part B* **2010,** *65*, 24-32.

23. Shaw, D. M.; Joensuu, O. I.; Ahrens, L. H. A double-arc method for spectrochemical analysis of geological materials. *Spectrochimica Acta* **1950,** *4*, 233-236.

24. King, R. B. "Arthur Scott King, Biographical Memoirs". *National Academy of Sciences* **2010**, viewed on 22 July 2010, http://www.nap.edu/readingroom.php?book=biomems&page=aking.html

25. Castro Sousa, J. K.; de Sousa Dantas, A. N.; Marques, A. L. B.; Lopes, G. S. Experimental design applied to the development of a copper direct determination method in gasoline samples by graphite furnace atomic absorption spectrometry. *Fuel Processing Technology* **2008,** *89*, 1180-1185.

26. Cassella, R. J.; de Sant'Ana, O. D.; Santelli, R. E. Determination of arsenic in petroleum refinery streams by electrothermal atomic absorption spectrometry after multivariate optimization based on Doehlert design. *Spectrochimica Acta Part B* **2002,** *57*, 1967-1978.

27. Micó, C.; Recatalá, L.; Peris, M.; Sánchez, J. Assessing heavy metals sources in agricultural soils of an European Mediterrean area by multivariate analysis. *Chemosphere* **2006,** *65*, 863-872.

28. Do, B.; Robinet, S.; Pradeau, D.; Guyon, F. Speciation of arsenic and selenium compounds by ion-pair reversed-phase chromatography with electrothermic atomic absorption spectrometry. Application of experimental design for chromatographic optimisation. *Journal of chromatography A* **2001,** *918*, 87-98.

29. Ojeda, C. B.; Rojas, F. S. Determination of rhodium: Since the origins until today. Atomic absorption spectrometry. *Talanta* **2005,** *68* (2006), 1407-1420.

30. Najafi, N. M.; Shahparvizi, S.; Rafati, H.; Ghasemi, E.; Alizadeh, R. Preconcentration and determination of ultra-traces of platinum in human serum using the combined electrodeposition-electrothermal atomic absorption spectroscopy (ED-ETAAS) and chemometric method. *Journal of Pharmaceutical and Biomedical Analysis* **2010,** *53*, 58-61.

31. Stoll, C. "Clifford Stoll Quotes". *BrainyQuote* **2010**, viewed on 22 July 2010, http://www.brainyquote.com/quotes/authors/c/clifford_stoll.html

32. Duckworth, J. H. Spectroscopic Quantitative Analysis. In *Applied Spectroscopy, A Compact Reference for Practitioners*, Workman, J., Springsteen A.W., Eds.; Academic Press: San Diego, 1998; p 93.

33. Wikipedia "Matryoshka doll". *Wikipedia* **2010**, viewed on 22 July 2010, http://en.wikipedia.org/wiki/Matryoshka

34. Geladi, P.; Esbensen, K. H. The start and early history of chemometrics: selected interviews. Part 1. *Journal of Chemometrics* **1990,** *4*, 337-354.

35. Esbensen, K. H.; Geladi, P. The start and early history of chemometrics: selected interviews. Part 2. *Journal of Chemometrics* **1990,** *4*, 389-412.

36. Kowalski, B. R. Chemometrics. *Analytical Chemistry* **1980,** *52*, 112R-122R.

37. Lavine, B. K. Chemometrics. *Analytical Chemistry* **2000,** *72*, 91R-97R.

38. Kiralj, R.; Ferreira, M. C. M. The past, present and future of chemometrics worldwide: some etymological, linguistic, and bibliometric investigations. *Journal of Chemometrics* **2006,** (20), 247-272.

39. Pretsch, E.; Wilkins, C. L. Use and abuse of chemometrics. *Trends in Analytical Chemistry* **2006,** *25* (11).

40. Lavine, B.; Workman, J. Chemometrics. *Analytical Chemistry* **2008,** *80*, 4519-4531.

41. Frank, I. E.; Kowalski, B. R. Chemometrics. *Analytical Chemistry* **1982,** *54*, 232R-243R.

42. Kowalski, B. R. Chemometrics: view and propositions. *Jornal of Chemical Information and Computer Sciences* **1975,** *15* (4), 201-203.

43. Geladi, P.; Esbensen, K. H. Chemometrics, a growing and maturing discipline. (Editorial). *Chemometrics and Intelligent laboratory Systems* **1990,** *7*, 197.

44. Delaney, M. F. Chemometrics. *Analytical Chemistry* **1984,** *56*, 261R-277R.

45. Workman, J. The state of multivariate thinking for scientists in industry: 1980-2000. *Chemometrics and Intelligent laboratory Systems* **2002,** *60*, 13-23.

46. Brown, S. D.; Blank, T. B.; Sum, S. T.; Weyer, L. G. Chemometrics. *Analytical Chemistry* **1994,** *66*, 315R-359R.

47. Ramos, L. S.; Beebe, K. R.; Carey, W. P.; Sanchez, E.; Erickson, B. C.; Wilson, B. E.; Wangen, L. E.; Kowalski, B. R. Chemometrics. *Analytical Chemistry* **1986,** *58*, 294R-315R.

48. Brown, S. D.; Sum, S. T.; Despagne, F. Chemometrics. *Analytical Chemistry* **1996,** *68*, 21R-61R.

49. Lavine, B.; Workman, J. Chemometrics. *Analytical Chemistry* **2004,** *76*, 3365-3372.

50. Lavine, B.; Workman, J. Chemometrics. *Analytical Chemistry* **2006,** *78*, 4137.

51. Pierce, K. M.; Hoggard, J. C.; Mohler, R. E.; Synovec, R. E. Recent advancements in comprehensive two-dimensional separations with chemometrics. *Journal of chromatography A* **2008,** *1184*, 341-352.

52. Aleme, H. G.; Costa L.M.; Barbeira P.J.S. Determination of gasoline origin by distillation curves and multivariate analysis. *Fuel* **2008,** *87*, 3664-3668.

53. van Mispelaar, V. G.; Smilde, A. K.; de Noord, O. E.; Blomberg, J.; Schoenmakers, P. J. Classification of highly similar crude oils using data sets from comprehensive two-dimensional gas chromatography and multivariate techniques. *Journal of chromatography A* **2005,** *1096*, 156-164.

54. Jos, A.; Moreno, I.; González, A. G.; Repetto, G.; Cameán, A. M. Differentiation of sparkling wines (cava and champagne) according to their mineral content. *Talanta* **2004,** *63*, 377-382.

55. Li, J.; He, M.; Han, W.; Gu, Y. Analysis and assessment on heavy metal sources in the coastal soils developed from alluvial deposits using multivariate statistical methods. *Journal of Hazardous Materials* **2009,** *164*, 976-981.

56. Felipe-Sotelo, M.; Andrade, J. M.; Prada, D. Partial Least Squeares Multivariate Regression as an alternative to handle interferences of Fe on the determination of trace Cr in water by electrothermal atomic absorption spectrometry. *Analytical Chemistry* **2003,** *75*, 5254-5261.

57. Tsalev, D. L.; Slaveykova, V. I. Chemical modifications in electrothermal atomic absorption spectrometry. Organization and classification of data by multivariate methods. *Journal of Analytical Atomic Spectrometry* **1992,** *7*, 147-153.

58. Vankeerberghen, P.; Smeyers-Verbeke, J. The quality coefficient as a tool in decisions about the quality of calibration in graphite furnace atomic absorption spectrometry. *Chemometrics and Intelligent laboratory Systems* **1992,** *15*, 195-202.

59. Grotti, M.; Magi, E.; Leardi, R. Study of interfences in graphite furnace atomic absorption spectrometry by means of experimental design. *Analytica Chimica Acta* **1996,** *327*, 47-51.

60. Grotti, M.; Leardi, R.; Frache, R. Empirical modelling of interferences in electrothermal atomization atomic absorption spectrometry. *Analytica Chimica Acta* **1998,** *376*, 293-304.

61. Grotti, M.; Leardi, R.; Gnecco, C.; Frache, R. Determination of manganese by graphite furnace atomic absorption spectrometry: matrix effect control by multiple linear regression model. *Spectrochimica Acta part B* **1999,** *54*, 845-851.

62. Grotti, M.; Abelmoschi, M. L.; Soggia, F.; Tiberiade, C.; Frache, R. Reduction of interferences in graphite furnace atomic absorption spectrometry by multiple linear regression modeling. *Spectrochimica Acta part B* **2000,** *55*, 1847-1860.

63. Esbensen, K. H. *Multivariate data analysis - in practice;* 5 ed.; CAMO: 2002.

64. Brereton, R. G. *Applied chemometrics for scientists;* Wiley: 2007.

65. Adams, M. J. *Chemometrics in analytical spectroscopy.;* 2 ed.; The Royal Society of Chemistry: 2004; Vol. 8.

66. Olivieri, A. C. Analytical advantages of multivariate data processing. One, two, three, infinity? *Analytical Chemistry* **2008,** *80*, 5713-5720.

67. Doyle, A. C. "Quotation by Sir Arthur Conan Doyle". *Dictionary. com* **2010**, viewed on 22 July 2010, http://quotes.dictionary.com/From_a_drop_of_water_a_logician_could

68. Lakein, A. "Alan Lakein quotes". *Thinkexist. com* **2010**, viewed on 2 July 2010, http://thinkexist.com/quotation/failing_to_plan_is_planning_to_fail/252276.html

69. Deming, S. N.; Morgan, S. L. Teaching the fundamentals of experimental design. *Analytica Chimica Acta* **1983,** *150*, 183-198.

70. Morgan, E.; Burton, K. W.; Church, P. A. Practical exploratory experimental design. *Chemometrics and Intelligent laboratory Systems* **1989,** *5*, 283-382.

71. Bruns, R. E.; Scarmino, I. S.; de Barros Neto, B. *Statistical design - Chemometrics;* Elsevier: 2006; Vol. 25.

72. Le Garrec, H.; Giamarchi, P.; Cabon, J.; Le Bihan, A. Electrothermic factors optimization in electrothermal atomic absorption spetrometry via an optimal experimental design matrix. *Analytica Chimica Acta* **1998,** *368*, 59-70.

73. de Amorim, F. R.; Bof, C.; Franco, M. B.; da Silva, J. B. B.; Nascentes, C. C. Comparative study of conventional and multivariate methods for aluminium determination in soft drinks by graphite furnace atomic absorption spectrometry. *Microchemical Journal* **2006,** *82*, 168-173.

74. Leardi, R. Experimental design in chemistry: A tutorial. *Analytica Chimica Acta* **2009,** *652* (1-2), 161-172.

75. Hintermaier, J. C. Foundations for experimental design. *Analytical Chemistry* **1948,** *20* (12), 1144-1146.

76. González, A. G. Two level factorial experimental design based on multiple linear regression models: a tutorial digest illustrated by case studies. *Analytica Chimica Acta* **1998,** *360*, 227-241.

77. Lobo, F. A.; Goveia, D.; de Oliveira, A. P.; Pereira-Filho, E. R.; Fraceto, L. F.; Dias Filho, N. L.; Rosa, A. H. Comparison of the univariate and multivariate methods in the

otpimiazation of experimental conditions for determining Cu, Pb, Ni and Cd in biodisiel by GFAAS. *Fuel* **2009,** *88*, 1907-1914.

78. Kettaneh-Wold, N. Use of experimental design in the pharmaceutical industry. *Journal of Pharmaceutical & Biomedical Analysis* **1991,** *9* (8), 605-610.

79. Araujo, P. W.; Brereton, R. G. Experimental design III. Quantification. *Trends in Analytical Chemistry* **1996,** *15* (3), 156-163.

80. Ellison, S. L. R.; Barwick, V. J.; Farrant, T. J. D. *Practical Statistics for the Analytical Scientist. A Bench Guide.;* 2 ed.; Royal Society of Chemistry: Cambridge, 2009.

81. Izgi, B.; Demir, C.; Güçer, A. Application of factorial design for mercury determination by trapping and graphite furnace atomic absorption spectrometry. *Spectrochimica Acta Part B* **2000,** *55*, 971-977.

82. Trindade, J.; Marques, A.; Lopes, G.; Marques, E.; Zhang, J. Arsenic determination in gasoline by hydride generation atomic absorption spectroscopy combined with a factorial experimental desigh approach. *Fuel* **2006,** 2155-2161.

83. Atkinson, A. C.; Tobias, R. D. Optimal experimental design in chromatography. *Journal of chromatography A* **2008,** *1177*, 1-11.

84. Araujo, P. W.; Brereton, R. G. Experimental Design I. Screening. *Trends in Analytical Chemistry* **1996,** *15* (1), 26-31.

85. Bermejo-Barrera, P.; Moreda-Piñero, A.; Muñiz-Naveiro, O.; Gomez-Fernandez, A. M. J.; Bermejo-Barrera, A. Optimization of a microwave-pseudo-digestion procedure by experimental design for the determination of trace elements in seafood products by atomic absorption spectrometry. *Spectrochimica Acta Part B* **2000,** *55*, 1351-1371.

86. Bermejo-Barrera, P.; Muñiz-Naveiro, O.; Moreda-Piñero, A.; Bermejo-Barrera, A. Experimental designs in the optimisation of ultrasonic bath-acid-leaching procedures for the determination of trace elements in human hair samples by atomic absorption spectrometry. *Forensic Science International* **2000,** *107*, 105-120.

87. Koch, I.; Harrington, C. F.; Reimer, K. J.; Cullen, W. R. Simplex optimisation of conditions for the determination of antimony in environmental samples by using electrothermal atomic absorption spectrometry. *Talanta* **1997,** *44*, 771-780.

88. El Ati-Hellal, M.; Hellal, F.; Dachraoui, M.; Hedhili, A. Plackett-Burman designs in the pretreatment of macroalgae for Pb, Cr and Al determinations by GFAAS. *C. R. Chimie* **2007,** *10*, 839-849.

89. Araujo, P. W.; Brereton, R. G. Experimental design II. Optimization. *Trends in Analytical Chemistry* **1996,** *15* (2), 63-70.

90. Ferreira, S. L. C.; dos Santos, H. C.; Fernandes, M. S.; de Carvalho, M. S. Application of Doehlert matrix and factorial designs in optimization of experimental variables associated with preconcentration and determination of molybdenum in sea-water by inductively coupled plasma optical emission spectrometry. *Journal of Analytical Atomic Spectrometry* **2002,** *17*, 115-120.

91. Brown, J. H.; Gomez, M. J.; Benzo, Z.; Vaz, J. E. Application of the response surface methodology for potassium determination in soils by AAS using the slurry technique. *Chemometrics and Intelligent laboratory Systems* **1996,** *35*, 239-247.

92. Einstein, A. "Albert Einstein Quotes". *BrainyQuote* **2010**, viewed on 22 July 2010, http://www.brainyquote.com/quotes/quotes/a/alberteins100017.html

93. Puling, L. "Dr Linus Pauling quotes". *Thinkexist. com* **2010**, viewed on 22 July 2010, http://thinkexist.com/quotation/facts_are_the_air_of_scientists-without_them_you/213998.html

94. *Design-Expert,* version 8.0.1; Stat-Ease, Inc: 2010

95. Campbell, W. C.; Ottaway, J. M. Atom-formation processes in carbon-furnace atomizers used in atomic-absorption spectrometry. *Talanta* **1974,** *21*, 837-844.

96. L'vov, B. V.; Nikolaev, V. G.; Novichikhin, A. V.; Polzik, L. K. Effect of platform material on sample vaporisation rate in graphite furnace atomic absorption spectrometry. *Spectrochimica Acta* **1988,** *43B* (911), 1141-1146.

97. Ortner, H. M.; Schlemmer, G.; Welz, B. Scanning electron microscopy studies on surfaces from electrothermal atomic absorption spectrometry - I. polycrystalline electrographite tubes with and without pyrographite coating. *Spectrochimica Acta* **1985,** *40B* (7), 959-977.

98. Welz, B.; Curtius, A. J.; Schlemmer, G. Scanning electron microscopy studies on surfaces from electrothermal atomic absorption spectrometry - III. The lanthanum modifier and the determination of phosphorus. *Spectrochimica Acta* **1986,** *41B* (11), 1175-1201.

99. Djojoputro, H.; Ismadji, S. Density and viscosity of several aldehydes fragrance compounds in their binary mixtures with ethanol at (298.15 K, 308.15 K, and 318.15 K). *Journal of Chemical and Engineering Data* **2010,** *50* (6), 2003-2007.

100. Haynes, W. M.; Lide, D. R. "CRC handbook of chemistry and physics. Internet Version.". *CRC Handbook of Chemistry and Physics* **2010**, viewed on 06 October 2010, http://www.hbcpnetbase.com/

101. Luebke, W. "Octanal (aldehyde C-8)". *The Good Scents Company* **2010**, viewed on 06 October 2010, http://www.thegoodscentscompany.com/data/rw1000251.html

102. Castillo, J. R.; Mir, J. M.; Bendicho, C. Effect of matrix components on chromium atomization processes in graphite furnace atomic absorption spectrometry. *Spectrochimica Acta* **1987,** *43B* (3), 263-271.

103. Sardans, J.; Montes, F.; Penuelas, J. Determination of As, Cd, Cu, Hg and Pb in biological samples by modern electrothermal atomic absorption spectrometry. *Spectrochimica Acta Part B* **2010,** *65*, 97-112.

APPENDIX A – ANOVA TABLES

This appendix lists the ANOVA tables of the results discussed in Chapter 4. The reader should note that the factorial designs did not comprise any replicate or centre point (with the exception of Tables A.10, A.11 and A.12).

Table A.1: ANOVA table generated by the 2^{7-4} FFD applied to the Methanol sample (Section 4.2).

Source	Sum of Squares	df	Mean Square	F-Value	P-Value
Model	0.00381725	2	0.001908625	12.54642564	0.0113
Rt_{Ash} (D)	0.000496125	1	0.000496125	3.261298274	0.1308
Rt_{Atom} (G)	0.003321125	1	0.003321125	21.831553	0.0055
Residual	0.000760625	5	0.000152125		
Cor Total	0.004577875	7			

Table A.2: ANOVA table generated by the 2^{7-4} FFD applied to the Octanal sample (Section 4.2).

Source	Sum of Squares	df	Mean Square	F-Value	P-Value
Model	0.0039485	2	0.00197425	11.20459705	0.0142
Rt_{Ash} (D)	0.0004205	1	0.0004205	2.386492622	0.1830
Rt_{Atom} (G)	0.003528	1	0.003528	20.02270148	0.0066
Residual	0.000881	5	0.0001762		
Cor Total	0.0048295	7			

Table A.3: ANOVA table generated by the 2^{7-4} FFD applied to the Toluene sample (Section 4.2).

Source	Sum of Squares	df	Mean Square	F-Value	P-Value
Model	0.004278125	1	0.004278125	23.90570431	0.0027
Rt_{Atom} (G)	0.004278125	1	0.004278125	23.90570431	0.0027
Residual	0.00107375	6	0.000178958		
Cor Total	0.005351875	7			

Table A.4: ANOVA table generated by the 2^{7-4} FFD applied to the Hexane sample (Section 4.2).

Source	Sum of Squares	df	Mean Square	F-Value	P-Value
Model	0.00282025	2	0.001410125	9.330851944	0.0205
T_{Tr} (C)	0.000990125	1	0.000990125	6.551695616	0.0507
Rt_{Atom} (G)	0.001830125	1	0.001830125	12.11000827	0.0177
Residual	0.000755625	5	0.000151125		
Cor Total	0.003575875	7			

Table A.5: ANOVA table generated by the 2^{6-3} FFD applied to the Mehanol sample (Section 4.3).

Source	Sum of Squares	df	Mean Square	F-Value	P-Value
Model	0.0002205	1	0.0002205	13.56923077	0.0103
T_{Ash} (F)	0.0002205	1	0.0002205	13.56923077	0.0103
Residual	9.75E-05	6	0.00001625		
Cor Total	0.000318	7			

Table A.6: ANOVA table generated by the 2^{6-3} FFD applied to the Octanal sample (Section 4.3).

Source	Sum of Squares	df	Mean Square	F-Value	P-Value
Model	0.0021451	1	0.00214513	6.792848661	0.0403
T_{Ash} (F)	0.0021451	1	0.00214513	6.792848661	0.0403
Residual	0.0018948	6	0.00031579		
Cor Total	0.0040399	7			

Table A.7: ANOVA table generated by the 2^{6-3} FFD applied to the Toluene sample (Section 4.3).

Source	Sum of Squares	df	Mean Square	F-Value	P-Value
Model	0.0005951	1	0.00059513	6.289299868	0.0460
T_{Ash} (F)	0.0005951	1	0.00059513	6.289299868	0.0460
Residual	0.0005678	6	9.4625E-05		
Cor Total	0.0011629	7			

Table A.8: ANOVA table generated by the 2^{6-3} FFD applied to the Hexane sample (Section 4.3).

Source	Sum of Squares	df	Mean Square	F-Value	P-Value
Model	0.00073	2	0.000365	5.738993711	0.0507
Rt_{Ash} (D)	0.000392	1	0.000392	6.163522013	0.0557
Ht_{Ash} (E)	0.000338	1	0.000338	5.314465409	0.0693
Residual	0.000318	5	0.0000636		
Cor Total	0.001048	7			

Table A.9: ANOVA table generated by the 2^{6-3} FFD applied to the Water sample (Section 4.3).

Source	Sum of Squares	df	Mean Square	F-Value	P-Value
Model	0.0026645	1	0.0026645	3.311651994	0.1187
T_{Ash} (F)	0.0026645	1	0.0026645	3.311651994	0.1187
Residual	0.0048275	6	0.00080458		
Cor Total	0.007492	7			

Table A.10: ANOVA table generated by the 2^2 CFD applied to the Methanol sample (Section 4.4).

Source	Sum of Squares	df	Mean Square	F-Value	P-Value
Model	0.0193	2	0.00965	19.37751004	0.0491
Rt_{Atom} (A)	0.000256	1	0.000256	0.514056225	0.5478
T_{Atom} (B)	0.019044	1	0.019044	38.24096386	0.0252
Residual	0.000996	2	0.000498		
Cor Total	0.020296	4			

Table A.11: ANOVA table generated by the 2^2 CFD applied to the Octanal sample (Section 4.4).

Source	Sum of Squares	df	Mean Square	F-Value	P-Value
Model	0.0093305	2	0.00466525	185.4970179	0.0054
Rt_{Atom} (A)	0.0002103	1	0.00021025	8.359840954	0.1017
T_{Atom} (B)	0.0091203	1	0.00912025	362.6341948	0.0027
Residual	5.03E-05	2	0.00002515		
Cor Total	0.0093808	4			

Table A.12: ANOVA table generated by the 2^2 CFD applied to the Toluene sample (Section 4.4).

Source	Sum of Squares	df	Mean Square	F-Value	P-Value
Model	0.0081005	2	0.00405025	344.7021277	0.0029
Rt_{Atom} (A)	9.025E-05	1	9.025E-05	7.680851064	0.1093
T_{Atom} (B)	0.0080103	1	0.00801025	681.7234043	0.0015
Residual	0.0000235	2	0.00001175		
Cor Total	0.008124	4			

APPENDIX B – CONFOUNDING CONSEQUENCES

This appendix shows a practical example of confounding. The results obtained from the 2^{7-4} FFD (Section 4.2) are computed after the deletion of the Rt_{Atom} factor and are expressed in terms of N-Plots, Pareto charts and ANOVA tables. These results can be compared to the results described in Section 4.2 to underline how the Rt_{Atom} factor, after its removal from the design, is replaced by the Rt_{Tr}-Rt_{Atom} interaction (AF), one of its *aliased* terms (Appendix C).

Each factor is identified with a capital letter: A-Rt_{Tr}; B-Ht_{Tr}; C-T_{Tr}; D-Rt_{Ash}; E-Ht_{Ash}; F-T_{Ash}.

B1 Methanol

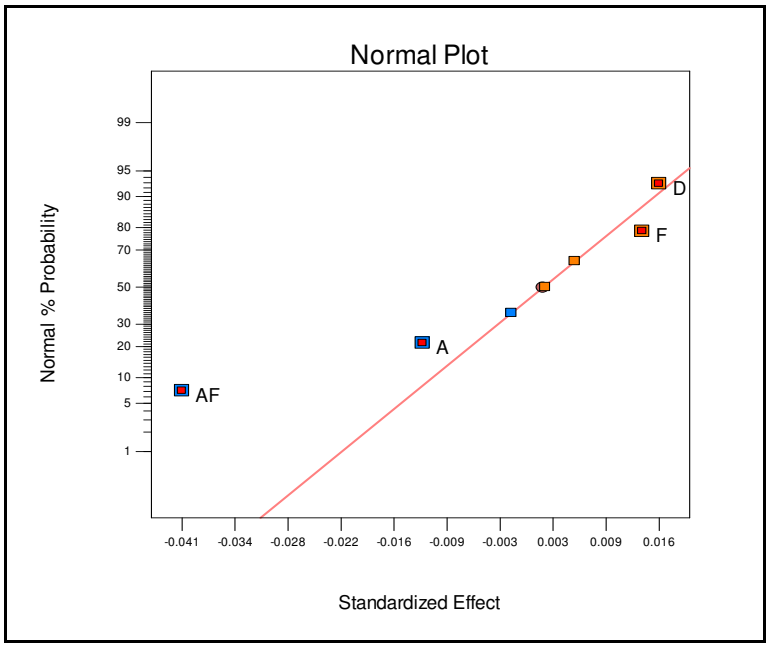

Figure B.1: N-Plot obtained after the deletion of the Rt_{Atom} factor (G) from the 2^{7-4} FFD applied to the Methanol sample.

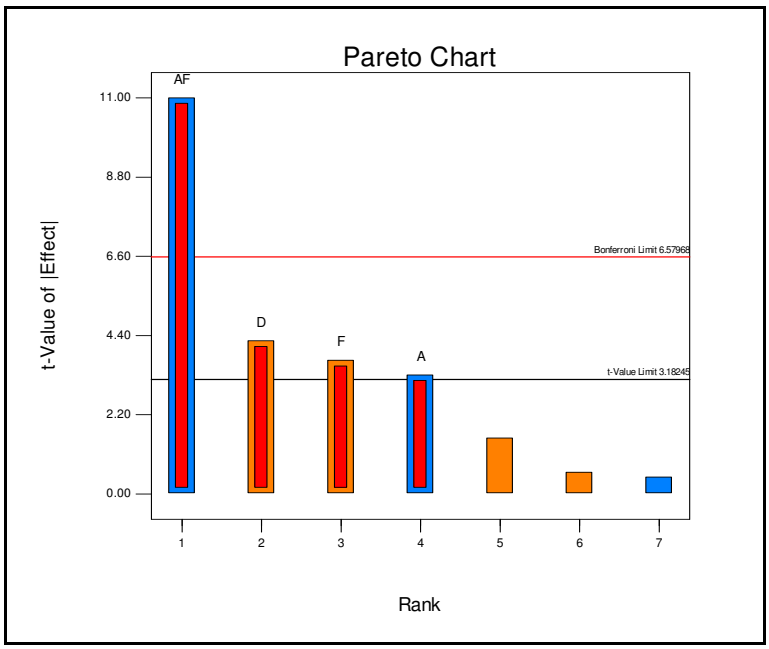

Figure B.2: Pareto chart obtained after the deletion of the Rt_{Atom} factor (G) from the 2^{7-4} FFD applied to the Methanol sample.

Table B.1: ANOVA table obtained after the deletion of the Rt_{Atom} factor (G) from the 2^{7-4} FFD applied to the Methanol sample.

Source	Sum of Squares	df	Mean Square	F-Value	P-Value
Model	0.004496	4	0.001123875	40.93019727	0.0060
Rt_{Tr} (A)	0.0003	1	0.000300125	10.93019727	0.0455
Rt_{Ash} (D)	0.000496	1	0.000496125	18.06828528	0.0239
T_{Ash} (F)	0.000378	1	0.000378125	13.77086495	0.0340
AF interaction	0.003321	1	0.003321125	120.9514416	0.0016
Residual	8.24E-05	3	2.74583E-05		
Cor Total	0.004578	7			

B2 Octanal

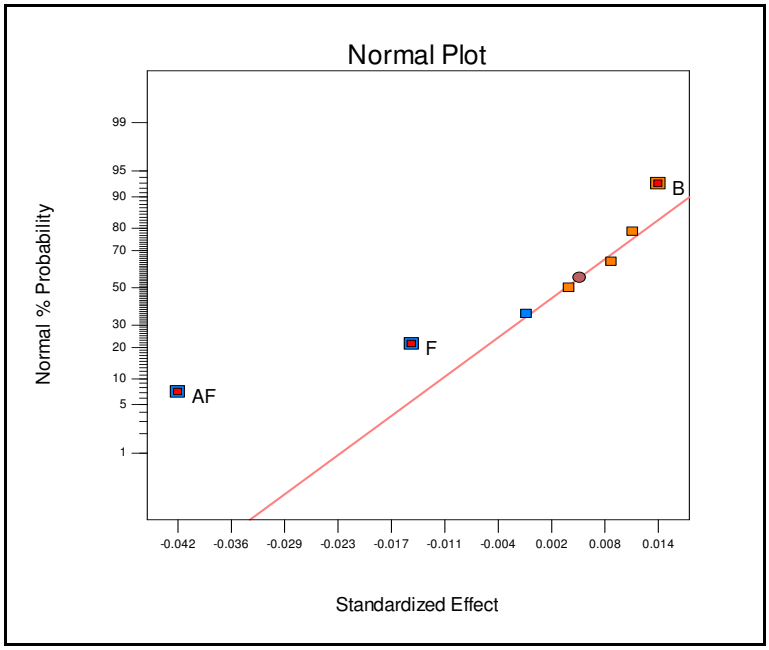

Figure B.3: N-Plot obtained after the deletion of the Rt_{Atom} factor (G) from the 2^{7-4} FFD applied to the Octanal sample.

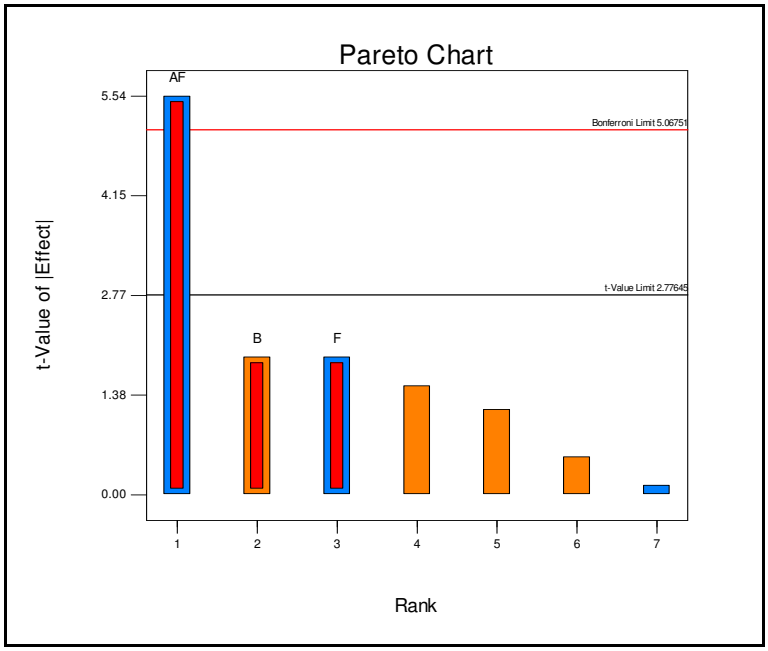

Figure B.4: Pareto chart obtained after the deletion of the Rt_{Atom} factor (G) from the 2^{7-4} FFD applied to the Octanal sample.

Table B.2: ANOVA table obtained after the deletion of the Rt_{Atom} factor (G) from the 2^{7-4} FFD applied to the Octanal sample.

Source	Sum of Squares	df	Mean Square	F-Value	P-Value
Model	0.004369	3	0.001456333	12.6500181	0.0165
Ht_{Tr} (A)	0.000421	1	0.0004205	3.652551574	0.1286
T_{Ash} (F)	0.000421	1	0.0004205	3.652551574	0.1286
AF interaction	0.003528	1	0.003528	30.64495114	0.0052
Residual	0.000461	4	0.000115125		
Cor Total	0.00483	7			

B3 Toluene

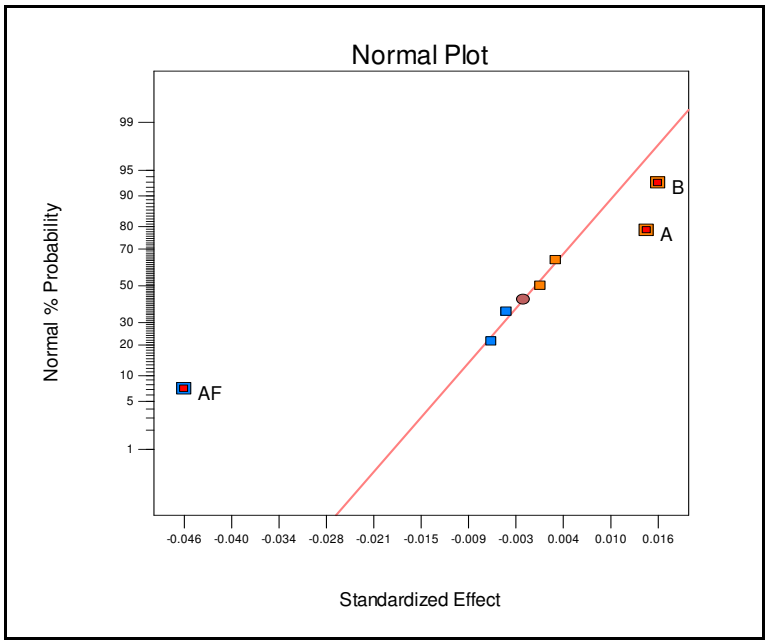

Figure B.5: N-Plot obtained after the deletion of the Rt_{Atom} factor (G) from the 2^{7-4} FFD applied to the Toluene sample.

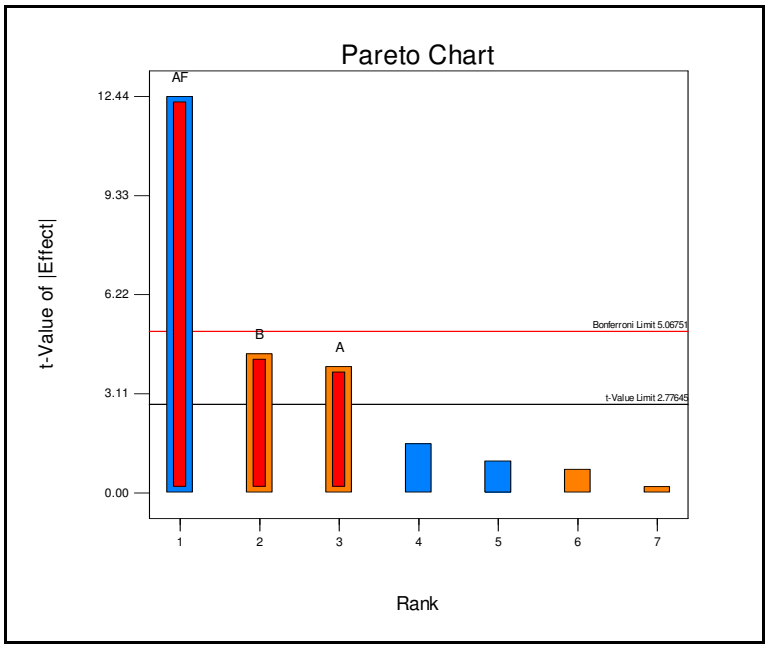

Figure B.6: Pareto chart obtained after the deletion of the Rt_{Atom} factor (G) from the 2^{7-4} FFD applied to the Toluene sample.

Table B.3: ANOVA table obtained after the deletion of the Rt_{Atom} factor (G) from the 2^{7-4} FFD applied to the Toluene sample.

Source	Sum of Squares	df	Mean Square	F-Value	P-Value
Model	0.005241	3	0.001747125	63.24434389	0.0008
Rt_{Tr} (A)	0.000435	1	0.000435125	15.75113122	0.0166
Ht_{Tr} (B)	0.000528	1	0.000528125	19.11764706	0.0119
AF interaction	0.004278	1	0.004278125	154.8642534	0.0002
Residual	0.000111	4	0.000027625		
Cor Total	0.005352	7			

B4 Hexane

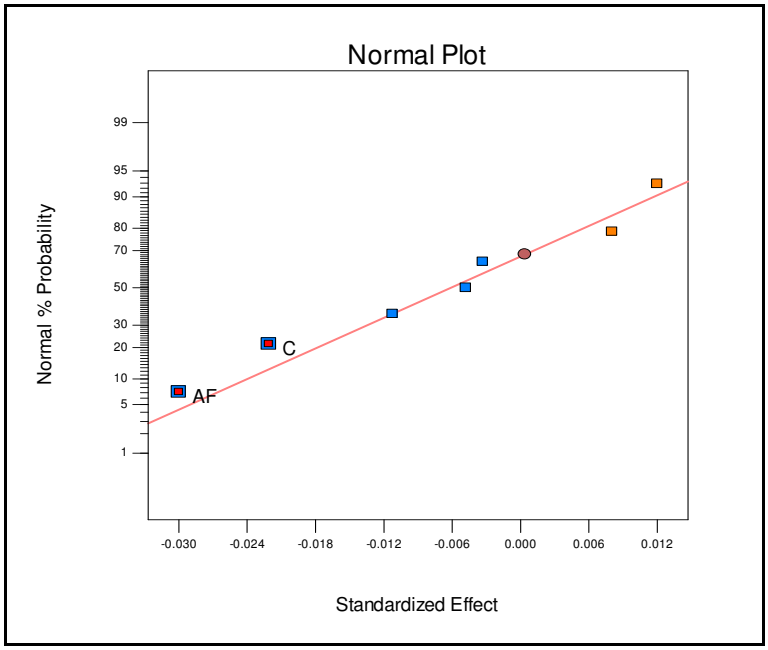

Figure B.7: N-Plot obtained after the deletion of the Rt_{Atom} factor (G) from the 2^{7-4} FFD applied to the Hexane sample.

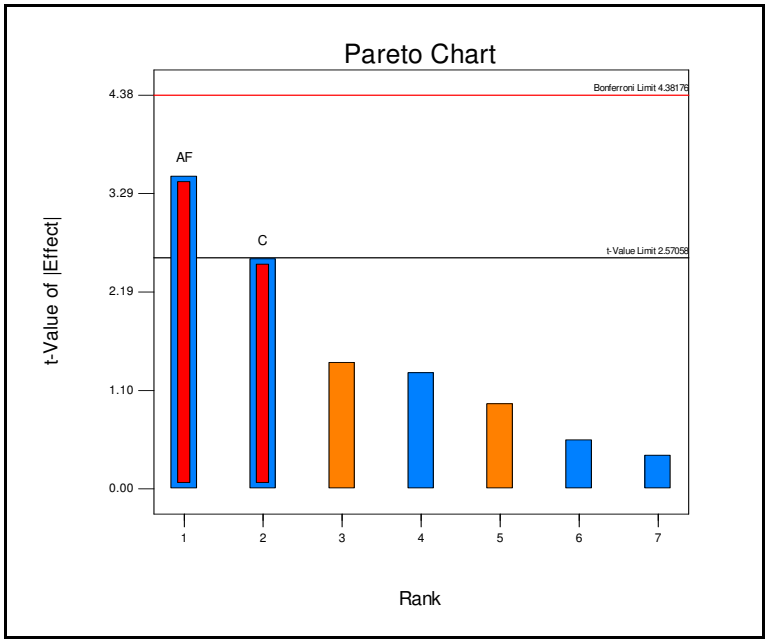

Figure B.8: Pareto chart obtained after the deletion of the Rt_{Atom} factor (G) from the 2^{7-4} FFD applied to the Hexane sample.

Table B.4: ANOVA table obtained after the deletion of the Rt_{Atom} factor (G) from the 2^{7-4} FFD applied to the Hexane sample.

Source	Sum of Squares	df	Mean Square	F-Value	P-Value
Model	0.00282	2	0.001410125	9.330851944	0.0205
T_{Tr} (C)	0.00099	1	0.000990125	6.551695616	0.0507
AF interaction	0.00183	1	0.001830125	12.11000827	0.0177
Residual	0.000756	5	0.000151125		
Cor Total	0.003576	7			

APPENDIX C – FFD ALIASED TERMS

Tables C.1 and C.2 list the aliased terms of the experimental designs discussed in Sections 4.2 and 4.3, respectively.

Table C.1: Aliased terms from the 2^{7-4} FFD.

Term	Aliases
Intercept	ABD ACE AFG BCF BEG CDG DEF
Rt_{Tr} (A)	BD CE FG BCG BEF CDF DEG
Ht_{Tr} (B)	AD CF EG ACG AEF CDE DFG
T_{Tr} (C)	AE BF DG ABG ADF BDE EFG
Rt_{Ash} (D)	AB CG EF ACF AEG BCE BFG
Ht_{Ash} (E)	AC BG DF ABF ADG BCD CFG
T_{Ash}. (F)	AG BC DE ABE ACD BDG CEG
Rt_{Atom} (G)	AF BE CD ABC ADE BDF CEF

Table C.2: Aliased terms from the 2^{6-3} FFD.

Term	Aliases
Intercept	ABD ACE BCF DEF
Rt_{Tr} (A)	BD CE BEF CDF
Ht_{Tr} (B)	AD CF AEF CDE
T_{Tr} (C)	AE BF ADF BDE
Rt_{Ash} (D)	AB EF ACF BCE
Ht_{Ash} (E)	AC DF ABF BCD
T_{Ash}. (F)	BC DE ABE ACD
AF	BE CD ABC ADE BDF CEF